Processes in Working Memory and their Impact on Long-term Memory Formation in Young and Older Adults

Cumulative Thesis

presented to the Faculty of Arts and Social Sciences

of the University of Zurich

for the degree of Doctor of Philosophy

by Lea Maria Bartsch

from Aachen, Germany

Accepted in the fall semester 2018

on the recommendation of the doctoral committee:

Prof. Dr. Klaus Oberauer

Prof. Dr. Yee Lee Shing

Zurich, 2018

Abstract

The goal of my dissertation project is to isolate processes people engage in during working memory (WM) encoding and maintenance, and to further examine the impact of these processes on long-term memory (LTM) formation. The interest arises from the hypothesis that individual and group differences in long-term memory abilities might reflect (age-related) variations in WM (Zacks, Hasher, & Li, 2000). Further, the present work was inspired by the assumption that maintenance of information in WM relies on cognitive processes such as refreshing in which attention is redirected to representations of the to-be-remembered information, and elaboration in which memory representations are enriched through links to related knowledge in LTM.

I conducted two lines of experiments to test the above assumptions: In the first I collect behavioural as well as neural evidence *for* a distinction of refreshing and elaboration. The behavioural data show that elaboration does not benefit WM but LTM of young adults. The latter is not the case for older adults, which hints at a potential cause for memory deficits related to aging. Further, I applied multivariate pattern analyses (MVPA) of functional magnetic resonance imaging (fMRI) data to identify and differentiate brain activation patterns associated with refreshing or elaborating memory items during a WM task in young and older adults. In the second line of experiments, I investigated the importance of establishing and maintaining bindings in WM to age differences in associative LTM. Equating WM for bindings between young and older adults reduced, but did not fully eliminate, the associative LTM deficit in the older adults. I conclude that reduced WM capacity does not cause the LTM deficit of older adults. Rather, both WM and LTM deficits are reflections of a common cause, that is related to time afforded for encoding. Taken together, both lines of experiments provide a deeper insight into the relationship of WM and LTM, with evidence for their distinct susceptibility to elaboration as well as evidence against the causal role of WM capacity limits to LTM deficits in older adults.

Zusammenfassung

Das Ziel meines Dissertationsprojekts ist es, Prozesse zu isolieren, die bei der Enkodierung und Aufrechterhaltung des Arbeitsgedächtnisses (AG) eine Rolle spielen, und die Auswirkungen dieser Prozesse auf die Langzeitgedächtnisbildung (LZG) zu untersuchen. Das Interesse folgt aus der Hypothese, dass (altersbedingte)Inter-individuelle- und Gruppenunterschiede bei der Langzeitgedächtnisbildung , Variationen im WM widerspiegeln (Zacks, Hasher & Li, 2000) Darüber hinaus wurde die vorliegende Arbeit von der Annahme motiviert, dass die Aufrechterhaltung von Informationen im AG auf kognitiven Prozessen beruht, wie zum Beispiel dem *Refreshing*, bei dem die Aufmerksamkeit auf Repräsentationen der zu erinnerenden Information fokussiert wird, und die *Elaboration*, bei der die Gedächtnisrepräsentationen durch Verbindungen zu verwandten LZG -Netzwerken angereichert werden.

Ich habe zwei Versuchsreihen durchgeführt, um die obigen Annahmen zu überprüfen: In der ersten zeige ich sowohl verhaltens- als auch neuronale Evidenz für eine Unterscheidung von *Refreshing* und *Elaboration*. Die Verhaltensdaten zeigen, dass die *Elaboration* nicht das AG, aber das LZG von jungen Erwachsenen stärkt. Letzteres ist bei älteren Erwachsenen nicht der Fall, was auf eine mögliche Ursache für altersbedingte Gedächtnisdefizite hindeutet. Weiterhin habe ich multivariate Muster-Analysen (MVPA) von funktionellen Magnetresonanztomographie (fMRI) - Daten implementiert, um Gehirnaktivierungsmuster zu identifizieren und zu differenzieren, die mit dem *Refreshing* oder der *Elaboration* von Gedächtnisinhalten während einer AG-Aufgabe bei jungen und älteren Erwachsenen assoziiert sind. In der zweiten Versuchsreihe untersuchte ich, wie wichtig es für das assoziativen LZG ist, Bindungen im AG zu etablieren und aufrechtzuerhalten. Ausgeglichenes AG für Bindungen zwischen jungen und älteren Erwachsenen reduzierte, aber eliminierte das assoziative LZG-Defizit in älteren Erwachsenen nicht vollständig. Ich komme zu dem Schluss, dass eine verringerte AG-Kapazität nicht das LZG-Defizit älterer Erwachsener verursacht. Vielmehr sind sowohl AG als auch LZG-Defizite Abbild einer gemeinsamen Ursache, die mit der für die Enkodierung zur Verfügung stehenden Zeit zusammenhängt. Zusammengefasst gewähren beide Experimentreihen einen tieferen Einblick in das Verhältnis von AG und LZG mit Hinweisen auf ihre ausgeprägte Anfälligkeit für die Ausarbeitung sowie Evidenz gegen die kausale Rolle von AG-Kapazitätsgrenzen für LZG-Defizite bei älteren Erwachsenen.

Contents

Part I – Synopsis

1. Introduction

Working memory (WM) is understood as a system for holding a limited amount of information available for processing (Baddeley, 1986), whereas episodic long-term memory (LTM) is conceptualized as the system for permanently storing, managing, and retrieving information for later use, and its capacity could be unlimited (Tulving, 1972). Two important aspects of these systems are currently under debate: (1) the architecture of the human memory system in general as well as (2) what limits WM capacity. Within the first debate one group of models understands WM and LTM as two separable systems of memory (Atkinson & Shiffrin, 1968; Baddeley, 2012; Barrouillet & Camos, 2015), whereas others conceptualize memory as a unitary system (e.g., Crowder, 1982; Melton, 1963; Nairne, 1990, 2002). Intermediate theories that fill positions in between these extreme views conceptualize WM as a subset of LTM representations that – for a limited time – are in a qualitatively higher state of accessibility (Cowan, 1995; McElree, 1998, 2006; Oberauer, 2002).

One of these intermediate conceptualizations is the Three Embedded Components Model (Oberauer, 2002, 2009) in which WM comprises of three functionally distinct sets of representations, embedded in LTM: 1) the activated part of LTM, 2) the region of direct access, and 3) the focus of attention. In the model, LTM is understood as an associated network of representations, which can be activated through perceptual input or spread of activation from associated representations. Those subsequently activated representations constitute the first component of WM, namely the activated part of LTM. The second component is the region of direct access: a subset of those activated representations is bound to contexts, establishing item-context bindings. These serve as retrieval cues and could for instance link the item to its serial

position in a list, its location, or to a feature such as an object's shape or color. Following from this, the function of the region of direct access is to allow for quick access to a subset of representations. The third component comprises of the focus of attention, which selects one[1] element among those currently held in the region of direct access for concurrent processing.

What follows from the model is a central role for relational information (i.e. associations & bindings) in the organization of memory representations: On the one hand, WM is conceptualized as a system built of item-context bindings, within the aforementioned region of direct access. It is responsible for their formation and to potentially use them as retrieval cues. On the other hand, LTM is understood as an associated network of representations. As far as this relational information (i.e. bindings/associations) is common to both – conceptualizations of WM and LTM – there are important distinctions.

Bindings in WM are in a sense arbitrary connections of content to other information, such as their serial position. The resulting relational information has an immediate significance to the system but not all of the concurrent input has to be stored necessarily into LTM. Further, and more importantly, the bindings in WM do not necessarily rely on semantic content, whereas this is what defines an associative network of representations in LTM. I therefore hypothesized that if the memory system of WM and LTM are as closely related as proposed above, WM bindings could be strengthened through a mechanism which is similarly beneficial to associative representational networks of LTM – namely through elaboration.

By semantically enriching bindings in WM (such as item-to-serial position) through elaboration, these would transform into associative structures in LTM, and thus benefit WM in

[1] The number of items held in the focus of attention is currently still under debate. See Oberauer and Bialkova (2009), Oberauer and (2012) and Gilchrist and Cowan (2011) for a discussion.

two ways: 1) Bindings being transferred into LTM would free capacity in WM and 2) semantically enriched bindings would serve as stronger retrieval cues, than arbitrary ones. If bindings in WM can be strengthened through processes that involve the semantic content, this would become evident in immediate as well as delayed memory performance. Furthermore, any deficiency related to the processing (formation, maintenance, strengthening) of bindings would have wide consequences - for instance it may be the cause of WM- as well as LTM-deficits in older adults.

Specifically, I investigated this main hypothesis in three different ways. (A1) I directly investigate the potential immediate and delayed behavioral benefits of strengthening the representations in WM through elaboration and refreshing and (A2) collect neural evidence for these processes' distinct nature and differential effects on WM and LTM. Further, (B) I investigate the role of a deficiency in bindings in WM to associative LTM.

(A) The potential benefits of strengthening representations

For a better understanding of aspect (A), let's revisit the following basic WM task: Participants are asked to encode and later recall a list of nouns in their serial order, while the inter-stimulus-interval (ISI) is experimentally manipulated. A common finding in this WM task is that the performance of recalling items in their serial order (immediately or after a delay) benefits from extended free time between the individual items at encoding (i.e. a longer ISI). The Three Embedded Components model accounts for this finding by assuming that the perceptual input elicited by the words activate their semantic representations in LTM. In the region of direct access, these active representations are bound to the nouns' serial position in the list. At test, a probe asking for the item presented at a certain serial position, then serves as a retrieval cue for

the specific noun. Following this logic, better recall arises from better retrieval cues, which in turn are the result of an increase in time to strengthen item-context bindings.

Indeed, the benefit of a prolonged ISI has been primarily attributed to rehearsal processes that act upon the memory representations during that free time (Camos & Portrat, 2015; Souza & Oberauer, 2017; Vergauwe, Camos, & Barrouillet, 2014; Vergauwe & Langerock, 2017). Yet, two points are under debate here: First, it is unclear what the target of those rehearsal processes is: Whether they strengthen the context-content bindings, or whether they act upon the activation of content representations themselves, to prevent them from decay-based memory loss (Barrouillet, Bernardin, & Camos, 2004; Barrouillet, Bernardin, Portrat, Vergauwe, & Camos, 2007). Second, it is unclear which rehearsal process causes the (greatest) benefit. Currently three processes are discussed in the literature: people could be (1) attending the to-be-remembered information by refreshing it, (2) they could be elaborating on it, or (3) perform articulatory rehearsal. As an extension of this, it is unclear which cognitive mechanisms underlie refreshing and elaboration and in how far they are separable processes.

Although rehearsal processes are the prominent contender for explaining the memory benefits of extended ISIs, an alternative possibility is that free time is used to consolidate each just encoded item into WM (Bayliss, Bogdanovs, & Jarrold, 2015; De Schrijver & Barrouillet, 2017; Ricker, 2015; Ricker & Cowan, 2014).

(B) Age-related deficits in building and maintaining representations in WM and their effects on LTM

The role of age-related deficits in building and maintaining representations such as bindings for related aspects of human memory is an extension of the above debates. The

efficiency of binding mechanisms, including the strengthening of bindings through rehearsal processes, can be a limiting factor for memory performance and may explain age variations in WM as well as LTM. In detail, older adults may no longer enrich the arbitrary bindings in WM as efficiently with semantic content, resulting in weaker retrieval cues or reduced freeing of WM capacity through the transfer of information to associative LTM structures compared to young adults.

Apart from directly manipulating the processes through experimental instruction, it is possible to investigate the role of strengthening bindings in WM to LTM, by quantitatively equating the WM for bindings between two populations that have been shown to initially differ in memory performance (i.e. young and older adults), and then measure whether the same quantity of remembered bindings in WM has led to an equated LTM for those associations. In other words, whether equating WM between young and older adults leads to qualitatively different representations in older adults, that allowed the built-up of LTM just as in young adults.

The present work

Taken together, the view on WM suggested by the Three Embedded Component model, inspires an interest in the processes that potentially strengthen content-context bindings in WM. These processes may serve as a communication mechanism between WM and LTM and can tell us more about the architecture of our memory system. Finally, it inspires the interest in whether any deficiency in binding processes cause age-related memory deficits. The goal of this dissertation is to isolate processes of WM maintenance by investigating their impact on WM as well as LTM, to explore their underlying mechanisms and thereby extending the research on how WM and LTM communicate and exchange information. Furthermore, this thesis investigates the

processes' role to age-related deficits in LTM as well as the role of bindings in WM in explaining these deficits.

Specifically, I conducted two lines of experiments and implemented behavioural and neural measures to test the hypothesis that individual and group differences in long-term memory abilities reflect (age-related) variations in WM (Zacks, Hasher, & Li, 2000). Further, I tested the assumption that maintenance of information in WM relies on cognitive processes such as *refreshing* in which attention is redirected to representations of the to-be-remembered information, and *elaboration* in which memory representations are explicitly linked with related knowledge in LTM. In the first line of experiments, Study 1 introduces a behavioural paradigm to experimentally investigate the impact of elaboration and refreshing on WM and LTM and to test the hypothesis that WM performance benefits from semantically enriching the arbitrary item-serial position bindings in WM through elaboration similarly to how it benefits LTM. I further present results from Study 2 where I use multivariate pattern analyses (MVPA) of fMRI data to identify and differentiate brain activation patterns associated with refreshing or elaborating memory items during this very paradigm in young and older adults. In the second line of experiments, in Study 3, I investigate the importance of establishing and maintaining bindings in WM to age differences in associative LTM.

2. Study 1 – The Effects of Refreshing and Elaboration on Working Memory Performance, and their Contributions to Long-term Memory Formation

2.1 Theoretical Background

Refreshing is understood as briefly thinking of a stimulus just after it is no longer physically present but while its representation is still active (Johnson, Reeder, Raye, & Mitchell, 2002). The process was introduced as a general attention- based mechanism (attentional refreshing; see Camos et al., 2018 for a review) and is assumed to be distinct from articulatory rehearsal. Refreshing was proposed as a mechanism for enhancing and prolonging the activation of WM representations, by increasing an item's specific level of activation at each refreshing step (Lemaire, Pageot, Plancher, & Portrat, 2017; Portrat & Lemaire, 2015). Others have proposed that refreshing compared to repeated reading of a word also causes increased LTM (Johnson, Mitchell, Raye, & Greene, 2004; Johnson, Reeder, Raye, & Mitchell, 2002; Raye, Johnson, Mitchell, Greene, & Johnson, 2005; Raye, Johnson, Mitchell, Reeder, & Greene, 2002).

It is unclear *what* exactly is being reactivated by refreshing, thereby leaving under debate the mechanism of its potential benefit for memory (e.g. under conditions of extended time): The current views strongly depend on the researchers' understanding of the relationship of WM and LTM as well as their beliefs of what forms a WM representation. Some researchers argue that refreshing strengthens the bindings between items and their contexts, thereby creating stronger retrieval-cues for WM as well as LTM – this being in line with the conceptualization of WM in the Three Embedded Components model (Lewandowsky & Farrell, 2008; Loaiza & Mccabe, 2012; Oberauer & Hein, 2012; Oberauer & Lewandowsky, 2011). They attribute the WM benefit

of longer ISIs to the use of this free time to strengthen the bindings between items and their serial position through processes like refreshing, resulting in stronger retrieval cues at test.

Others understand refreshing as a mechanism for preserving and reconstructing the complex and multifaceted mental representations built in WM, which otherwise would suffer from temporal decay (Barrouillet & Camos, 2015; Johnson, 1992; Johnson & Johnson, 2009). Here the finding of WM benefitting from longer encoding times are explained the notion that more time to reactivate the items' representation counteracts decay more. In comparison, when less time is available more activation of memory traces is lost.

Lastly, some views assume that by focusing attention on an item during refreshing, it is more deeply encoded into memory and therefore better remembered, as predicted by the levels of processing theory (Craik & Lockhart, 1972). If that is the case, the effect of refreshing would be closely related to that of elaboration, the second proposed maintenance process in WM and the center of the main hypothesis. A deeper encoding into memory could also be conceptualized as forming stronger bindings not only to the items' context at encoding, but also enriching these arbitrary bindings with semantic content and as spreading activations in the associative network that forms our LTM to closely related representations, and thereby creating additional and more distinct retrieval cues.

2.2 Summary of Study 1

The goal of the first study was to isolate WM maintenance processes and to better understand their underlying mechanisms (see also Section 8). Further, the study aimed at experimentally testing the hypothesized role of refreshing and elaboration for WM and well as LTM formation. Finally, it was to investigate whether refreshing and elaboration are distinct

processes, or whether the process of focusing attention on a representation in WM (i.e. refreshing) stimulates the linking of the representation to existing networks in LTM (= elaboration). I addressed these goals by a series of two experiments in which subjects had to sequentially encode a list of six nouns in their serial order. Subsequently either the first or the last three-items had to be processed again according to one of four processing conditions: the items were either repeated with or without the instructions to form a vivid mental image of the items interacting or the items had to be refreshed – again with or without the elaboration instruction. In this case the symbols appeared as refreshing-cues in the locations of the to-be-refreshed items. This was followed by 4-alternatives forced choice task in which for each of the 6 serial positions the subjects had to choose from a set of possibilities including the correct, two lures and a new response option. The second experiment added two baseline conditions to assess the level of performance without any processing instruction: In the short condition, recognition followed immediately after encoding and in the long condition the participants were given "free time" for the same amount of time of the processing phase of the instructed conditions (6 seconds).

The experiments revealed evidence that, compared to a no-processing baseline, immediate memory was improved by repeating the items, but not by refreshing them. We replicated the long-term memory benefit for elaboration, but the results did not support its beneficial role for working memory. The baseline-comparison showed that refreshing preserves immediate memory but does not improve it beyond the level achieved without any processing. Further, there was no long-term memory benefit for refreshed compared to repeated items, which had been previously reported in the literature.

The results of Study 1 provide first evidence for a differentiation of refreshing and elaboration, as the latter showed a benefit for LTM performance, whereas refreshing did not. Interestingly, WM was not affected by elaboration, hinting at it not being a beneficial or necessary process for WM. Nevertheless, its benefits become evident after a delay, replicating the advantage of linking newly encoded information more deeply to its context for LTM. The exact mechanisms of the form of elaboration implemented here cannot be decisively uncovered by our paradigm, nevertheless it incorporates possible important components of such a mechanism. These include (1) the activation of semantic representations of the respective nouns as well as (2) the spread activation to related concepts, (3) the mental visualization and thereby creation of an integrated, newly constructed representation of the triplet, as well as (4) the creation of bindings between the three items. All of these proposed components would have resulted in memory benefits through the creation of qualitatively and quantitatively superior retrieval cues compared to repeated reading (as operationalized in the paradigm by repeated presentation without elaboration). Which of these possible mechanisms underlies the elaboration benefit in LTM, has to be the directly investigated in future studies.

3. Study 2 – Dissociating refreshing and elaboration by their neural signatures and their effects on working memory and long-term memory in young and old adults

3.1 Theoretical Background

First evidence for a distinction of refreshing and elaboration was provided in Study 1, where – based on behavioral results of young adults – a differential effect of the processes on LTM was shown. The results on WM were uninformative regarding the question of a functional distinction of refreshing and elaboration, which lead to the motivation to measure not only the

behavioral effects but also the processes elicited neural activation patterns via fMRI. This way, I wanted to investigate whether the processes of refreshing and elaboration are distinguishable during WM maintenance on a neural level and in how far the neural separability relates to each processes' behavioral consequences.

These processes have been the focus of several fMRI studies in the past, which either investigated elaboration or refreshing, but which independently found similar neural correlates. Further, subsequent memory effects in those studies provide evidence for the functional relation of the processes' elicited activation and the formation of memory traces, being tested either immediately or after a delay (for details see Section 9). In general, current models of the neurocognitive architecture of WM show consensus that the system is characterized by the interaction of frontal and posterior cortical areas, as well as sub-cortical structures (D'Esposito & Postle, 2015; Eriksson, Vogel, Lansner, Bergström, & Nyberg, 2015). More precisely, maintenance in WM supposedly results from the interactions of basic cognitive processes – such as attention – with representations in LTM, whereby prefrontally mediated control processes operate on posterior neural systems that are specialized for processing the particular perceptual representations. Important for my dissertation is the neural evidence incorporated in this consensus which supports the involvement of cognitive processes interacting as communication systems between new perceptual input and LTM representations. As discussed above, refreshing and elaboration have been advanced to serve this function.

Building on the hypothesized important role of rehearsal process for the establishment of representations and bindings in WM as well as their role for an integration of new information in LTM, any deficiency in these processes could cause for instance age-related memory deficits. Previous fMRI studies investigating refreshing and elaboration have provided evidence for this

claim, associating age-related changes in LTM performance with changes in process-dependent brain activations (Johnson et al., 2004). Furthermore, older adults appear to engage in memory-relevant strategic processes less efficiently than younger adults, which also speaks for the role of efficiency of rehearsal process contributing to age-related deficits in LTM (Dunlosky & Hertzog, 1998; Dunlosky, Hertzog, & Powell-Moman, 2005; Kausler, 1994). Regarding age-related changes in the neurocognitive architecture of WM, underactivations (e.g., Grady et al., 1998; Jonides et al., 2000; Rypma & D'Esposito, 2000) and overactivations (e.g., Cabeza, Anderson, Locantore, & McIntosh, 2002; Reuter-Lorenz et al., 2000) in older compared to younger adults have been interpreted as a signature of less differentiated representations and processing circuits (Baltes & Lindenberger, 1997; Li, Lindenberger, & Sikström, 2001). Given those ambiguous findings, the implementation of MVPA methods in order to analyse the representational content promises a new approach to investigate whether cognitive processes such as refreshing and elaboration are also distinguishable in older adults during the maintenance phase of a WM task and in order to also relate neural separability measures to the processes' behavioural outcomes. This will help to answer the question whether the efficiency of rehearsal processes (neurally and behaviorally) contributes to age-related deficit in LTM.

3.2 Summary of Study 2

As the consensus further hints towards a broad involvement of regions, it made sense to not investigate whether refreshing and elaboration are "located" in different regions of the brain, but rather more importantly in this case, whether the activation patterns within the regions of the fronto-parietal circuit associated with WM show distinguishable representational content between conditions of refreshing and elaboration. Furthermore, to assess the processes'

individual neural footprint and these measures' relation to the processes behavioral outcome, it was important to compare the individual processes to a process that is atypical for maintenance in WM, namely the re-encoding the perceptual information via repeated reading. The goal of Study 2 was to isolate the processes of refreshing and elaboration based on their behavioral outcome, their neural separability as assessed by the classifiability of the machine learning algorithm via MVPA and finally based on their role for age-related memory deficits. Included in this was also the goal to better understand the mechanism underlying the memory benefit of refreshing, relating to the third view above, whether by refreshing items in WM, they are more deeply encoded, similar to the process of elaboration.

In a combined mask of a priori brain regions from frontal, temporal, and parietal lobes, we found successful differentiation of brain activity associated with all three processes: *repeating*, *refreshing* and *elaboration.* Critically, the degree of neural separation between these processes within an individual was predictive of their memory performance. Re-reading items benefited WM performance more than refreshing did, but this relative advantage was reduced when the neural processes of reading and refreshing were more similar. That is, refreshing benefited WM when it appeared, in the brain, to be like repeated reading. Elaboration produced no benefit to WM, but did improve LTM, and this benefit increased as the neural separation between elaboration and repeating increased. Importantly, we were able to replicate the neural differentiation of these three processes in a sample of 27 older adults. In contrast to younger adults, we found that elaboration despite its similar neural separability, did not benefit LTM. This is a first hint, that the LTM deficits in older adults could be caused by a deficiency in semantically enriching or more deeply encoding the memoranda in WM.

4. Study 3 – Does limited working-memory capacity underlie age differences in associative long-term memory?

4.1 Theoretical Background

Bindings are proposed to play an important role in the architecture of human memory. Models like the Three Embedded Component model conceptualize WM as the system for building and maintaining bindings and further by proposing also LTM as a network of associated representations (Oberauer, 2002). For both WM and LTM the idea of connections between representations, their contexts, and related other representations are substantial. Following the above two studies on the role of strengthening representations and their associated context (i.e. serial position and other words of the triplet) through rehearsal processes, it was the goal of the third study to investigate whether any deficiency in the system's ability to form bindings causes other phenomena, like the pronounced age-related decline in associative memory (see Section 9). Although age can be seen as a critical source of inter-individual differences in various parts of memory performance, such as associative memory and also WM, it has no explanatory value and it is of central interest to unfold these age-related deficits by identifying mechanisms that both influence memory and vary with age.

I present the role of maintenance processes on item-context (i.e. item-serial position) bindings and their variation with age in Section 2 and Section 9, but these bindings are not the only ones playing a role in memory: observations of age-related associative memory deficits encompass different types of bindings, including also those between two items (Old & Naveh-Benjamin, 2008). Although the binding deficits seem to be overly present, recent findings suggest that older adults show a smaller deficit in item-context compared to item-item bindings (Overman, McCormick-Huhn, Dennis, Salerno, & Giglio, 2018). Extending the unfolding of

age-related memory deficits of Study 2, in which the paradigm focused on item-context (serial position) bindings, the third study investigated memory for pairs of words.

4.2 Summary of Study 3

The goal of Study 3 was to investigate whether the associative memory deficit in older adults is caused by binding deficits in WM (see also Section 9). I addressed this goal by conducting two experiments in which memory for word pairs was tested in samples of young and older adults. In Experiment 3.1, WM for bindings was equated between the age groups through an adaptive algorithm, which varied the presentation rate of word pairs based on ongoing binding performance. Critically, after memory was equated at the level of WM, associative LTM for those pairs was tested. If age-related differences in WM capacity truly cause the associative-memory deficit in LTM (as proposed by a *WM binding deficit account*), then equating WM binding performance between young and older adults should eliminate the age-related deficit in LTM. Alternatively, a *common cause account* would suggest that WM and LTM age deficits may be reflections of a common cause, which is partly compensated for by longer encoding time, leading to reductions in LTM binding deficits. For instance, both forms of memory might suffer from a similar age-related slowing of consolidation, the hypothetical process converting fragile, transient representations into more stable memory representations (Chun & Potter, 1995; Jolicœur & Dell'Acqua, 1998; Wixted, 2004). Therefore, I aimed to test the same hypothesis through a second approach in Experiment 3.2: If WM capacity limits the acquisition of associations in LTM, then increasing the load on WM (i.e., the number of presented pairs) should impair memory for bindings in WM as well as LTM.

Experiment 3.1 revealed evidence that, the successful equation of WM for bindings between age groups by adapting the presentation rate of to-be-remembered word pairs did *not* eliminate the LTM deficit in old age. This first evidence for the claim that older adults' LTM deficit is not entirely caused by a WM deficit at encoding, was further supported by Experiment 3.2. Despite its detrimental effect on both age groups' WM for bindings, increases in set size had no such effect on LTM for either age group. Conclusively the results of Study 3 have shown that what matters for LTM binding deficits in older adults is how long they attend to and process the individual pairs. Both WM and LTM deficits are better considered as reflections of a common cause that is related to time afforded for encoding.

5. General Discussion

The goal of the dissertation was to examine whether WM benefits from maintenance processes such as refreshing and elaboration, for example through semantically enriching the arbitrary item-context bindings of a word to its serial position, in line with these processes' hypothesized benefit for LTM. Following the proposed close relationship of WM and LTM as well as from the similarity between associative semantic networks and bindings in WM I predicted a similar benefit for remembering items in their serial order also for WM. Contrary to this prediction, studies 1 and 2 revealed that WM does not benefit from semantically enriching the item-context bindings through elaboration. Behavioral as well as neural evidence of study 1 and 2 reveal that LTM for items and their contexts (i.e. their list position relative to the other items in the list) benefit from elaboration, and that the process of elaboration was distinguishable from repeated reading in the brain.

Why is this the case? The first explanation entails that the comparison of repeated reading to repeated reading with the instruction to form a mental image is in so far not optimal, as WM (much less than LTM) already benefits from repeated reading, because it allows to simply re-encode the memoranda – reducing the need for active maintenance processes altogether. A closer look at previously reported elaboration benefits in LTM reveals that the effect is characterized by very low performance in comparison conditions of rote repetition (e.g. elaboration: 67% vs. rote repetition: 40% correct recognition (Davachi, Maril, & Wagner, 2001) or 66% vs. 28% correct recognition (Baker, Sanders, Maccotta, & Buckner, 2001)). In our study performance in the comparison condition was much higher for WM – due to the reasons stated above. Even though performance in WM for the condition of repeated reading was that high, observing an effect of elaboration on WM was theoretically still possible, as performance was far from ceiling (highest mean percentage correct $\approx$ 82%), leaving room for further improvement of subjects' WM through elaboration. Also, the neural data of study 2 show that subjects were indeed engaging differently in the processes of repeated reading compared to elaboration, and that this degree of the processes' neural separability related to the LTM benefit – with larger neural separability leading to larger LTM benefits – whereas it did not relate to WM.

Ensuing this, a second explanation could be that item-context bindings in WM and associations in LTM are of different nature and that elaboration is a mechanism optimized for the LTM entity of associative networks, but not for bindings in WM. In line with the claim that LTM associations and WM bindings are dissimilar entities, early work on patients with hippocampal damage suggested that the structure mediates LTM – but not WM – for relational information (Cohen & Eichenbaum, 1993). Nevertheless, more recent work has provided growing evidence that the hippocampus contributes to successful encoding and retrieval of relational information in

WM after all (Hannula & Ranganath, 2008) and that maintenance of multiple items in WM is accompanied by cross-frequency coupling of oscillatory activity in the hippocampus (Axmacher et al., 2010). This evidence speaks for the conceptualization of LTM associations and bindings in WM as very similar entities.

The third explanation acknowledges this last point by incorporating that associations in LTM and bindings in WM must not be fundamentally different, but that at different time-frames and with slightly varying testing procedures (immediate vs. delayed test) the cognitive system makes use of different retrieval cues. In the paradigm of study 1 and 2 the immediate test cues with the serial position, prompting the use of item-context bindings. In the delayed task, the participants are presented with the first and fourth items of a list and are asked to recognize other items of this list based on these item-cues – thereby stimulating the use of item-item bindings. Elaboration might specifically enrich the item-item bindings, allowing for the benefit of enriching these to become evident only in LTM and not WM.

Further, the fourth explanation for a lack of elaboration benefits in WM entails that WM and LTM are two systems which rely differently on semantic content: LTM strongly relies on semantic content, as shown by the behavioral LTM benefit of elaboration in young adults, which was further related to the differential engagement in the process in the brain. In WM however, where the maintained information itself is hypothesized to be constituted of the activated part of LTM (semantic) networks of associations (Oberauer, 2002), bindings of those activated representations to contexts, such as serial positions, do not benefit from additional semantic content (see Study 1 and 2). It seems like the WM system either does not use the additional information provided through semantically enriching item-context bindings, or that the transition from arbitrary item-context bindings to semantically enriched associations in LTM – which

would free WM capacity or serve as stronger retrieval cues – takes (longer) time. In the case of study 1 and 2, the delay between encoding and retrieval might not have sufficed for the effect to show in the immediate test. The results of study 3 are in line with this last point, providing evidence that deficiencies in LTM are better explained by a common cause which affects both WM and LTM, which is related to the available time at encoding, rather than by a WM capacity limit.

Taken together, these results speak against the initially proposed hypothesis – that WM bindings could be strengthened through a mechanism which is similarly beneficial to associative representational networks of LTM - and further speak against conceptualizations of memory as a unitary system (e.g. Crowder, 1982). That is, because the cognitive process of semantically enriching arbitrary bindings in WM did not have the same benefit on WM performance as it had on LTM. If WM and LTM were simply redundant labels for one and the same system, the engagement in a specific cognitive process should benefit both, WM and LTM and not only LTM.

The second part of this thesis' main hypothesis entailed an aging perspective, by proposing that due to the proposed close relationship of WM and LTM, and due the similarity of bindings to associative structures in LTM, any deficiency related to the processing (formation, maintenance, strengthening) of bindings in WM could cause WM- as well as LTM-deficits in older adults. The results of study 2 show that older adults were not able to benefit from enriching item-context bindings in WM efficiently with semantic content, neither for their immediate nor LTM performance. This finding suggests that although brain activation patterns revealed the engagement of a process different from mere reading, older adults were deficient in either using the semantic content to strengthen their memory, or in using these as better retrieval cues. The

former may have resulted from a deficiency to generate enriched representations in the first place or from the need for longer time to consolidate these. As discussed in section 8.2.3, previous work has pointed out the deficiency of older adults in producing elaborations. Two recent studies coincide with the view that if supported, older adults' LTM can benefit from semantic content: their associative LTM was higher when provided with schema-based information and it can be improved to the level of younger adults' when the association between a product and its price is similar to their real-world experience (Amer, Giovanello, Grady, & Hasher, 2018; Fine, Shing, & Naveh-Benjamin, 2018).

Study 3 further supports the second point from above, namely the notion that older adults' deficiency in WM for bindings as well as for associations in LTM is related to the time available for consolidation or other processes at encoding. Similarly, longer time for consolidating enriched item-context associations could have resulted in a LTM benefit in Study 2 also in older adults.

6. Future Directions

As discussed above, there are several possibilities that could explain the lack of an elaboration benefit for WM and new experiments are needed in order to decisively differentiate between these alternatives. Although it would be beneficial to further contribute to a mechanistic description of the process of elaboration, the evidence in this thesis suggest, that elaboration is not the process that people engage in during extended free time and which results in benefit for WM. Several instructions including refreshing and elaboration, have been investigated, but none reached the level of a no-instruction baseline. So far, it seems most beneficial to WM if people

don’t follow any instructions but instead, follow their own strategies, that probably also vary from trial to trial during this free time.

I propose the following approach as a possible future direction from the above results, to differentiate whether these processes are actually not the cause for the WM benefit of prolonged time or whether following the instructions per se hurt WM performance: I propose to collect fMRI data from young adults while performing two tasks. First, I would implement a localizer task for collecting an independent dataset on which a classifier would be trained to discriminate brain activation patterns associated with a number of candidate maintenance processes, including also trials without a specific instruction.

In the main experiment, a number of items would be presented, followed by a delay without any instructions and a recognition test. Using MVPA analyses methods, I would apply the classifier trained on distinguishing the candidate processes in the localizer task to the delay period of the main experiment, in order to then retrieve the evidence values for each of the above candidate processes. With this, I could get insight into what process people are engaging in during the maintenance delay. Further, the degree of engagement in each of the processes (i.e. the evidence values) could then be correlated to the WM outcome measures (recognition accuracy or reaction times), giving insight in the effectiveness of the processes.

A further follow-up of this dissertation lies in the aging data: I have isolated elaboration as a candidate process which is performed less successful in older adults, resulting in a LTM deficit compared to younger adults. Future studies should determine whether the time to engage in the process was insufficient and therefore did not lead to a LTM benefit in older adults or whether the difficulty for the older adults was to actually generate elaborations. An

operationalization could include a longer processing time for older adults, crossed with the factor of self-generated vs. provided elaborations.

7. Conclusion

The studies of this dissertation have provided evidence that processes of elaboration and refreshing are distinct in their behavioral as well as neural effects. Furthermore, instructed engagement in these processes does not benefit WM for items and their bindings to serial positions. The benefit of semantically enriching these relations through elaboration becomes evident only after a delay in LTM, whereas refreshing has no such effect. Older adults show a deficient use of elaborative processing, which given the results of the third study may be related to their need for more time to engage in the process. Taken together, these findings emphasize that although several candidate processes that are proposed to contribute to maintenance in WM have been isolated and experimentally investigated repeatedly, the processes that are initiated spontaneously without experimental control are more effective and seem to be different from the processes proposed so far.

Part II: Empirical Studies

8. The Effects of Refreshing and Elaboration on Working Memory Performance, and their Contributions to Long-term Memory Formation

Lea M. Bartsch, Henrik Singmann & Klaus Oberauer

University of Zurich

Status: Published in Memory & Cognition (July 2018)

Authors' contributions

LB reviewed the literature, designed the study, programmed the task with Matlab, organized data collection, analyzed the data and interpreted the results and wrote the manuscript

HS supported the analysis of the data and revised the manuscript

KO designed the research question, supervised and discussed LB's contributions, revised the manuscript

Author Note

We thank Raffael Schmitt and Samuel Pawel for helping with data collection. The data and the analysis scripts can be accessed in the Open Science Framework (https://osf.io/weuc2).

7.1 Abstract

Refreshing and elaboration are cognitive processes assumed to underlie verbal working memory maintenance and assumed to support long-term memory formation. Whereas refreshing refers to the attentional focusing on representations, elaboration refers to linking representations in working memory into existing semantic networks. We measured the impact of instructed refreshing and elaboration on working and long-term memory separately, and investigated to what extent both processes are distinct in their contributions to working as well as long-term memory. Compared to a no-processing baseline, immediate memory was improved by repeating the items, but not by refreshing them. There was no credible effect of elaboration on working memory, except when items were repeated at the same time. Long-term memory benefited from elaboration but not from refreshing the words. The results replicate the long-term memory benefit for elaboration, but do not support its beneficial role for working memory. Further, refreshing preserves immediate memory but does not improve it beyond the level achieved without any processing.

8.2 Introduction

The literature on memory has made a distinction between working memory and long-term memory for a long time (see Cowan, 2008, for a review). Working memory is understood as a system for holding a limited amount of information available for processing (Baddeley, 1986), whereas long-term memory is a system for permanently storing, managing, and retrieving information for later use with a probably unlimited capacity (Tulving, 1972). Theorists have often assumed one or several control processes that people could apply to the current contents of

working memory, which are thought to help maintaining information in working memory, establish the information in long-term memory, or both. Three such control processes are being discussed: People could (a) attend to the to-be-remembered information to refresh it; (b) elaborate on it; or (c) engage in articulatory rehearsal. Our study focusses on the experimental manipulation of two of these processes, namely refreshing and elaboration, and investigates their effects on (1) immediate memory (presumably reflecting maintenance in working memory) and (2) on delayed memory (reflecting episodic long-term memory).

8.2.1 Refreshing as maintenance mechanism in working memory

Refreshing is understood as briefly thinking of a stimulus just after it is no longer physically present but while its representation is still active (Johnson, Reeder, Raye, & Mitchell, 2002). The process was introduced as a general attention-based mechanism (attentional refreshing; Barrouillet & Camos, 2007; Cowan, 1999; Johnson, 1992) and is assumed to be distinct from articulatory rehearsal, which is conceptualized as the specialized mechanism for the verbal domain (Baddeley, 1986; Camos, Lagner, & Barrouillet, 2009; Hudjetz & Oberauer, 2007). Refreshing is a core component of several models of working memory: In the MEM (multiple-entry, modular) memory model, refreshing permits the reactivation of memory representations (Johnson, 1992). Similarly, Cowan (1995) states that a memory trace could be reactivated by focusing attention on it (again), before its activation is entirely lost. The time-based resource-sharing (TBRS) model proposes refreshing through attentional focusing as a mechanism for reactivating decaying memory traces (Barrouillet et al., 2004; 2007). Refreshing as conceptualized in the above theories has never been observed directly, but rather has been inferred from results of experiments varying the opportunity for refreshing (i.e., varying cognitive load, Barrouillet, Bernardin, Portrat, Vergauwe, & Camos, 2007; Camos, Mora, &

Barrouillet, 2013; Mora & Camos, 2013). The evidence for the presumed effects of refreshing is therefore less than compelling, as it does not rely on experimentally inducing the process in question. To date, only two studies (Souza et al., 2015; Souza & Oberauer, 2017) experimentally induced refreshing in a (visual) working memory task to test its effect on memory. Our study aimed at closing this gap for verbal material, by experimentally manipulating refreshing and investigating its role for working memory.

8.2.2 The role of refreshing for episodic long-term memory

Beside its supposed role in working memory maintenance, refreshing of information in working memory has also been argued to improve long-term memory (Johnson et al., 2002). Several studies have contrasted refreshing of a single word to repeated reading of a word, and consistently found a benefit of refreshing on delayed item recognition (Johnson et al., 2002; Johnson, Mitchell, Raye, & Greene, 2004; Raye, Johnson, Mitchell, Greene, & Johnson, 2005; Raye, Johnson, Mitchell, Reeder, & Greene, 2002).

Additional, more indirect evidence for the effect of refreshing on episodic memory comes from experiments varying the available time for refreshing. For instance, a study varying cognitive load in a complex-span task found that low cognitive load – providing more free time during the maintenance interval – led to better delayed recall (Camos & Portrat, 2015). The authors attributed this to the necessary involvement of refreshing during working memory maintenance to build up long-term memory representations. Converging evidence comes from the McCabe effect (McCabe, 2008): Words studied in complex span tasks are recalled better in a delayed memory test than words studied in simple span tasks. This effect is often explained by the hypothesis that the secondary task forces people to refresh items after each distraction, thereby generating better episodic retrieval-cues (Loaiza, Duperreault, Rhodes, & McCabe,

2015; Loaiza, Rhodes, & Anglin, 2013; Loaiza & McCabe, 2012a, 2012b). Contrary to this explanation, a recent study by Souza and Oberauer (2017b) showed that the long-term memory benefit of items in complex span tasks can be fully attributed to the longer amount of free time for processing information in working memory in complex compared to standard simple span tasks.

What do people do when asked to attend to a word just encoded into WM, or when given free time to process words during maintenance? One possibility is that by focusing attention on an item, it is more deeply encoded into memory and therefore better remembered, as predicted by the levels of processing theory (Craik & Lockhart, 1972). If that is the case, the effect of refreshing would be closely related to that of elaboration.

8.2.3 Elaboration effects on long-term memory

Elaboration refers to processes that more deeply encode and store information for later retrieval (*elaborative rehearsal;* Craik and Lockhart 1972; Greene, 1987; Klatsky, 1988). Elaboration is thought to enrich the memory representation of an item by activating many aspects of its meaning, and by linking it into the pre-existing network of semantic associations (Craik & Tulving, 1975). It has repeatedly been shown to improve episodic long-term memory (e.g. Craik & Tulving, 1975; Gallo, Meadow, Johnson, & Foster, 2008). Research has further focussed on the use of various elaborative strategies, such as mental imagery, sentence generation, or chunking, and provided evidence for their long-term memory benefits (e.g. Dunlosky & Hertzog, 2001).

8.2.4 Elaboration and working memory

Only little research has focussed on the effects of elaboration on immediate memory, and the results are inconclusive. Several studies (Bailey, Dunlosky, & Kane, 2008; Bailey, Dunlosky,

& Kane, 2011; Dunlosky & Kane, 2007; Kaakinen & Hyönä, 2007) investigated the effect of strategy use on performance on complex span tasks, a popular paradigm for testing working memory, and revealed a positive correlation between elaborative strategies and performance. In contrast, Morrison, Rosenbaum and colleagues (2016) found no correlation between semantic strategies and performance in a working memory task. However, this evidence is merely correlational. The only experimental evidence for a beneficial effect of elaboration for working memory has been shown in a very specific case by Jonker and Macleod (2015). In their study, they showed that an orienting task inducing relational processing of words during encoding resulted in equivalent memory for the serial order of a study list compared to silent reading. In contrast, any other orienting task to be performed on the items, including semantic judgement of the individual items, resulted in disruption of memory for order. Therefore, relative to other cognitive operations performed on the current contents of working memory, relational processing appears to help maintenance.

8.2.5 Are refreshing and elaboration distinguishable?

Conceptually, refreshing and elaboration are different processes, but it is not clear that their effects on memory are actually separable. It is conceivable that focusing attention on a representation in working memory leads to deeper and richer encoding, so that the effect of refreshing on memory is mediated by elaboration. If that is the case, refreshing and elaboration should have equivalent effects on both working memory and long-term memory. To date, no research has been reported on the relationship of refreshing and elaboration; with the present study, we aimed at providing some initial insight into their relation.

Taken together, the evidence for the beneficial effects of refreshing on working memory, so far, is only indirect, and the evidence for elaboration benefits on WM is only correlational. In

the following experiments, we aimed at closing that gap and experimentally controlled refreshing and elaboration through instruction. We investigated the effects of each process, and their combination, on an immediate and a delayed test of memory and investigated whether both processes show similar result patterns.

8.2.6 Design of Experiments

As in the studies of Johnson and colleagues, we compared instructed refreshing to a repeating (re-reading) baseline during the maintenance phase of a working-memory task. In two additional conditions we instructed participants to elaborate a subset of the items they held in memory. Elaboration logically entails attending to the words, either in memory or in the environment. When elaboration is applied to words just encoded into working memory, but no longer presented, it entails refreshing, whereas when elaboration is applied to words while they are presented, it entails (re-)reading, as in the repeat condition. Therefore, we realized two elaboration conditions: One in which words are repeated and elaborated, and one in which they are refreshed and elaborated. In this way we can gauge the effects of elaboration on its own by comparing elaboration of repeated words to the repeating baseline. In addition, we can ask whether combining elaboration with refreshing is more effective than each of them alone.

How can we measure the effect of refreshing in our paradigm? The Johnson et al. studies – testing the effect of refreshing on EM – used repeat as the baseline, and therefore we follow their precedent for assessing the effect of refreshing on EM. For assessing the effect of refreshing on WM, the repeat condition is probably not a suitable baseline because it provides a second chance for encoding the word into WM. Therefore, we assess the effect of refreshing against two baselines: The one used in Souza et al. (2016; i.e., comparison within the memory set between items refreshed more vs. less) and a comparison of refreshed items to a no-processing baseline.

For the first comparison, we compare the items that were processed in refreshing trials to the items within the same trial that were not further processed after initial encoding. The second comparison, against a no-processing baseline, is important to assess whether refreshing actually leads to an improvement of memory after encoding. This cannot be assessed with the first comparison.

8.3 Methods

In the two experiments presented here, we asked participants to remember six nouns in serial order. After list presentation, either the first three words or the last three words were to be processed again in one of four ways, depending on the experimental condition. During encoding it was not predictable which half of the items would have to be processed. In the repeat-without-elaboration condition, the three words appeared again sequentially on the screen, and the subjects had to simply re-read them silently. In the refresh-without-elaboration condition, the to-be-processed words were replaced by refreshing prompts appearing at the same location. The subjects were instructed to "think of" the corresponding words as soon as the prompts were shown. In the repeat-with-elaboration condition, the three to-be-processed words were shown again sequentially on the screen, and subjects were instructed to generate a vivid mental image of the three objects interacting. The stimuli appearing on the screen in that condition did not differ from the repeat-without-elaboration condition, leaving the instruction to form a vivid mental image as the only difference between these conditions. Finally, in the combined refresh-with-elaboration condition the participants had to "think of" the words replaced by the prompts, and additionally form a vivid mental image of those items. Again, the event sequence of this condition does not differ from the refresh-without-elaboration condition apart from the instruction to form a mental image. The experiments used a 2 x 2 x 2 (repeat/refresh [repeat,

refresh] x elaboration [with elaboration, without elaboration] x processing [processed triplet, not-processed triplet]) within-subject design. Orienting the processing task to only a subset of the words in memory allows us to draw inferences about the effect of each of the processing conditions on memory by comparing the memory performance of the triplet of words that was not further processed after encoding to the triplet that was processed according to one of the four experimental conditions.

8.4 Experiment 1

8.4.1 Participants

We recruited twenty students from the University of Zurich as participants (10 female) at the age of 19 – 28 years (M = 22.11 years SD = 5.32). They were compensated with either 15 Swiss Francs (about 15 USD) or partial course credit for the one-hour experiment.

1.2.1 Materials and procedure

The stimuli were nouns randomly drawn from a pool of 863 German abstract and concrete nouns for each subject. The nouns were between two and 15 letters long and had a mean normalised lemma frequency of 30.81/million (drawn from the dlexdb.de lexical database).

The sequence of an experimental trial is illustrated in Figure 1. The six to-be-remembered words in each trial were sequentially presented in boxes from top to bottom on the screen, each for 500 ms. Depending on the experimental condition a cue was presented 1000 ms after the last memory item, indicating whether the first half or the second half of the list had to be processed again. In the repeat-with-elaboration and repeat-without-elaboration conditions, the to-be-processed first word of the cued triplet was shown again in the same box as during encoding for 1400 ms, followed by a 600 ms inter-stimulus-interval (ISI). The respective 2nd and 3rd word of

that triplet followed at the same pace successively in their corresponding boxes. In the refresh-without-elaboration and refresh-with-elaboration conditions, each to-be-processed word of a triplet was replaced by a refreshing prompt (*#?#?#*) in its corresponding box, and participants were instructed to "think of" the word in that box. In the repeat-with-elaboration and refresh-with-elaboration conditions, participants were additionally instructed to form a vivid mental image of the three words interacting with each other.[2]

After processing the words in the cued triplet, participants' memory for each list item was tested in their order of presentation using a 4-alternative forced-choice procedure. For that purpose, four words were presented from which the subject could choose the correct word in the currently tested list position with a button press. All tests trials had the following four response options: the target (i.e., correct) word, one lure from the same triplet of words within the present list, one lure from the other triplet of the present list, and one new word. This choice had to be made for each of the serial positions successively. We applied this 4-alternatives forced choice recognition task in order to test both memory for items (i.e., discriminating between items that have been presented in the current memory list and new items) and for serial order (i.e., discriminating between the item in the tested position and other list items).

Within each block of four trials, the same type of processing was instructed throughout, and a screen repeating the instructions of the particular condition was shown prior to the beginning of each block. The order of the condition blocks was randomized between subjects. After a total of 16 blocks, with four blocks of each condition, the participants performed a mental-arithmetic task for two minutes, where they had to judge the correctness of multiplication statements (e.g. 3 x 6 = 18). After this distractor task, we assessed participants' long-term

memory for the words they had encoded into working memory throughout the experiment. To this end, we presented in each trial the first word of one word triplet in one memory list, and asked participants to choose the word that had followed the given word within the same triplet from four different options. These included the correct word (i.e., which could be either the word in the 2nd or 3rd position of the target triplet for the first prompt and the 5th or 6th word for the second prompt), a word from the other triplet of the same list, a word from another list, and a new word. This format allowed us to keep the format of delayed recognition very similar to the immediate test, and furthermore to compare in each trial the memory performance for processed to not-processed items. As for the immediate test, the delayed test provided information about both item memory (i.e., which words have been presented in the experiment) and relational memory (i.e., which words have been together in a triplet). The participants were made aware of the delayed memory test before the start of the experiment.

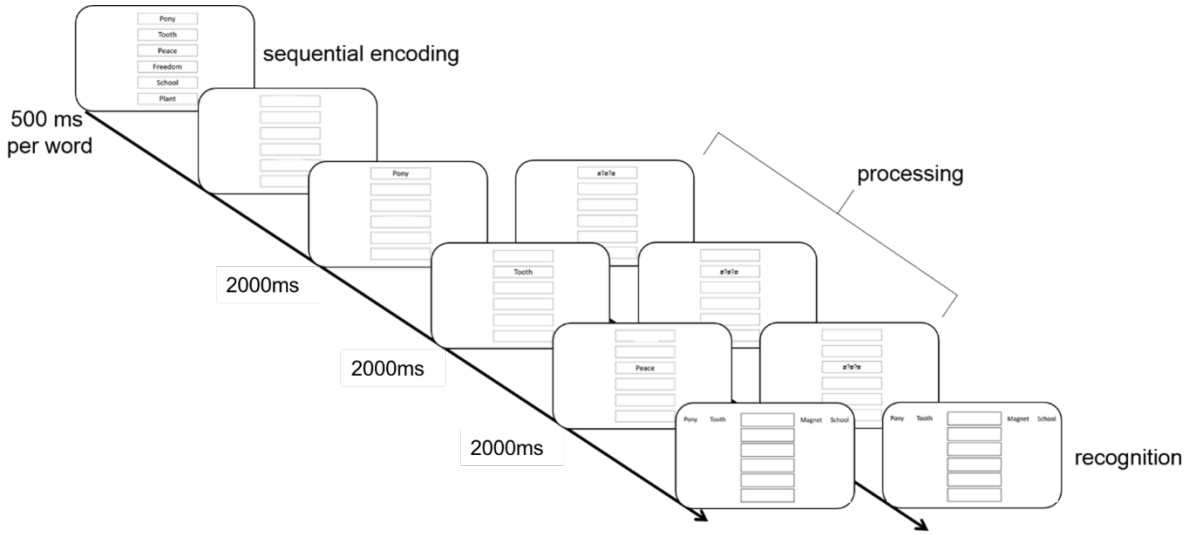


Figure 1. Illustration of the working memory paradigm of Experiment 1. Subjects were shown a list of six words sequentially, followed by either the first or second triplet being processed according to the four experimental conditions.

8.5 Experiment 2

Experiment 2 was identical to Experiment 1, but added two further conditions without any instructed processing. These conditions served as baselines against which we could gauge whether processing a subset of a memory list (according to the instructed process) improves or impairs memory for the processed triplet, and memory for the not-processed triplet, relative to a standard test of immediate memory. In the short baseline condition, recognition followed directly 1000 ms after the initial presentation of the list of six words, and allowed us to measure the level of memory directly after encoding. In the long baseline, a blank screen interval was inserted after list presentation for the same amount of time (six seconds) that the processing steps in the four experimental conditions took, before participants' memory was tested. The long baseline condition allowed us to investigate the impact of time on immediate memory without any processing manipulation, so participants were free to use it for any process on the memory items they might find helpful or do nothing.

8.5.1 Participants

For Experiment 2 we recruited 30 students from Zurich University (21 female) at the age of 19 – 28 years (M = 23.82 years SD = 3.82). They were compensated with either 15 Swiss Francs (about 15 USD) or partial course credit for the one-hour experiment.

8.5.2 Materials and procedures

The six conditions (the four experimental conditions of Experiment 1 and the two baseline conditions) were implemented within condition-pure mini-blocks of four trials, resulting in 12 trials per condition throughout the whole experiment. Everything apart from adding the two baseline conditions and reducing the number of trials per condition was held constant between Experiment 1 and 2.

8.6 Results

To draw inferences about the effect of refreshing and elaboration on working memory as well as these processes' impact on long-term memory formation, we first focus on the results of the four processing conditions of Experiments 1 and 2. We subsequently evaluate these results in comparison to the baseline conditions of Experiment 2. All data and analysis scripts can be assessed on the Open Science Framework (https://osf.io/weuc2).

8.6.1 Data Analysis

We analysed Experiments 1 and 2 jointly using a Bayesian generalized linear mixed model (BGLMM) implemented in the R package *rstanarm* (Stan Development Team, 2018). The dependent variable was the number of correct and incorrect responses in each cell of the design per participant. Correct responses were defined as choosing the target item from the four alternatives. Therefore, we assumed a binomial data distribution predicted by a linear model through a probit link function (i.e., a repeated-measures probit regression). The fixed-effects were processing (processed versus not-processed triplet), repeat/refresh (repeated versus refreshed items), elaboration (with versus without elaboration instruction), and all their interactions. Following the recommendation of Barr and colleagues (Barr, Levy, Scheepers, & Tily, 2013; see also Schielzeth & Forstmeier, 2009) we implemented the maximal random-effects structure justified by the design; by-participant random-intercepts and by-participant random-slopes for all fixed-effects (as all factors were within-subject factors). In addition, we estimated the correlation among the random-effects parameters. As factor coding we used the orthonormal contrasts described in Rouder et al. (2012; section 7.2) that guarantee that priors affect all factor levels equally. For factors with two levels as employed here this corresponds to contrasts with values of $\sqrt{2}/2$ and $-\sqrt{2}/2$.

Following Gelman et al. (2013), the regression coefficients were given weakly informative Cauchy priors with location 0 and scale 5. We used completely non-informative priors for the correlation matrices, so-called LKJ priors with shape parameter 1 (Stan Development Team, 2017). Bayesian procedures provide posterior probability distributions of the model parameters (i.e., the regression weights) that express uncertainty about the estimated parameters. The highest density regions (HDRs) of these posteriors can be used for statistical inference. A 95% HDR represents the range in which the true value of a parameter lies with probability 0.95, given model and data (Morey, Hoekstra, Rouder, Lee, & Wagenmakers, 2016). If zero lies outside the Bayesian HDR there is strong evidence for the existence of the corresponding effect; although the strength of evidence varies continuously, for simplicity we will describe effects as "credible" if their HDRs exclude zero. We used an MCMC algorithm (implemented in Stan; Carpenter et al., 2017) that estimated the posteriors by sampling parameter values proportional to the product of prior and likelihood. These samples are generated through 4 independent Markov chains, with 1000 warmup samples each, followed by 1000 samples drawn from the posterior distribution which were retained for analysis. Following Gelman and colleagues (2013), we confirmed that the 4 chains converged to the same posterior distribution by verifying that the $\hat{R}$ statistic – reflecting the ratio of between-chain variance to within-chain variance – was 1.02 for all parameters, and we visually inspected the chains for convergence.

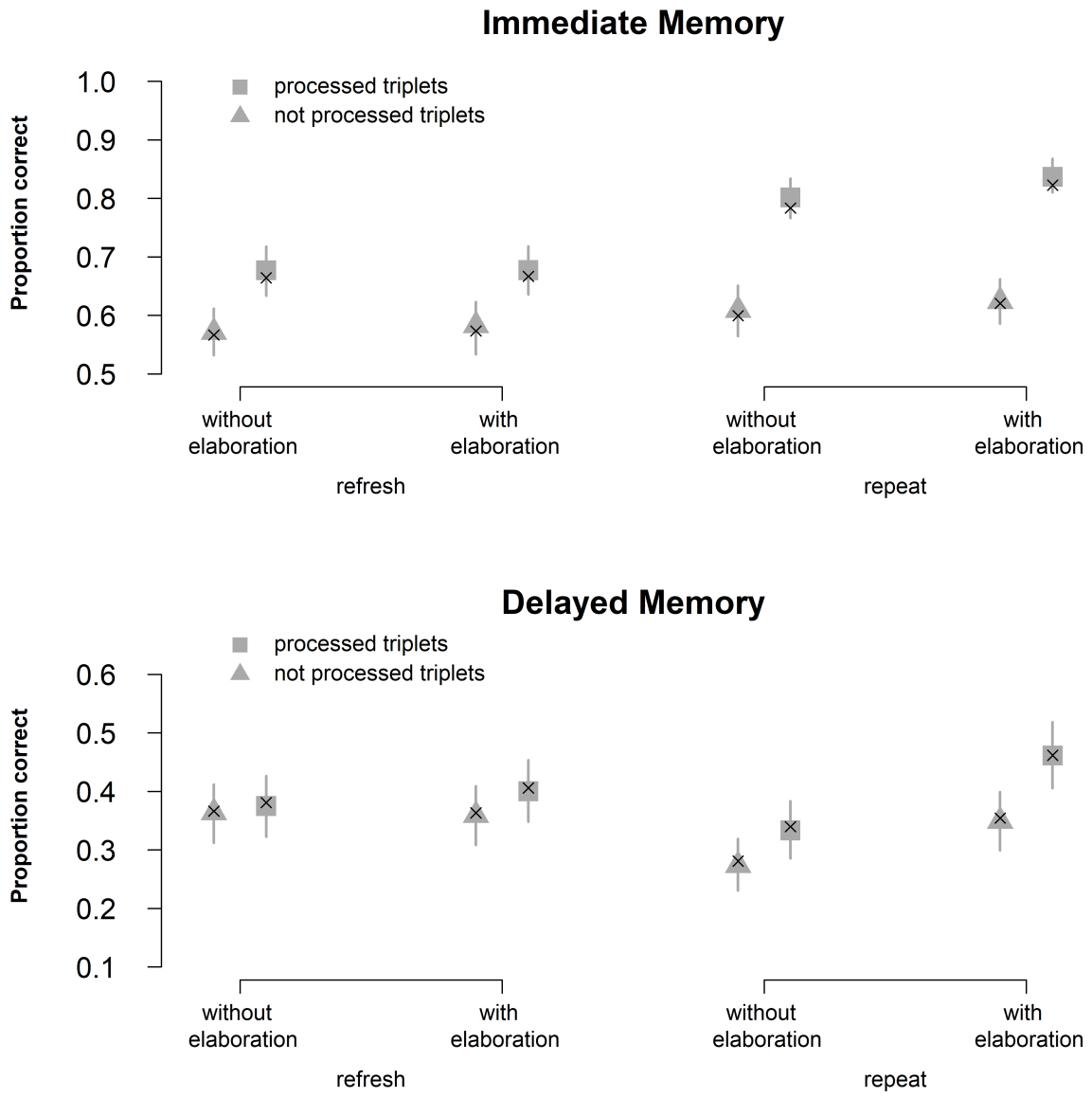


Figure 2 Proportion correct of the combined data from Experiments 1 and 2 in the immediate (WM, upper graph) and delayed (LTM, lower graph) memory task. The grey symbols and error bars represent estimated proportions and their 95% HDRs from the BGLMM. The crosses represent the observed proportions. Their overlap indicates that the model adequately describes the data.

For analysing the baseline conditions in Experiment 2, we estimated a second binomial BGLMM with two fixed-effects factors and their interaction on the same dependent variable. We again estimated the maximal random-effects structure and the correlation among the random effects parameters. The first factor was processing with two levels: processed or long baseline

versus not-processed or short baseline. This factor groups the conditions in which processing of words during maintenance is instructed or at least enabled, and the conditions providing no opportunity for such processing. The second factor combined repeat/refresh, elaboration, and baseline into a single grouping factor with five levels: repeat with elaboration, repeat without elaboration, refresh with elaboration, refresh without elaboration, and baseline. Note that we will focus here on the pairwise comparisons of the baseline conditions against the other factor levels. To test these pairwise comparisons we calculated difference distributions (i.e., posterior distributions of the differences of parameter values between two conditions) and report their 95% HDRs. Again, if this HDR does not contain zero this constitutes evidence for a highly credible difference. Again, we verified that the $\hat{R}$ statistic was 1.01, and visually inspected the chains for convergence.

8.6.2 Results

Figure 2 shows the estimated proportion of correct responses and their corresponding 95% highest posterior density regions for the immediate and delayed memory data from the core design shared by both experiments. The posterior effect estimates are presented in Table 1 and Table 2. As the BGLMM for the experimental conditions of Experiment 1 and 2 revealed evidence for the same pattern of results, we combined the data in a single BGLMM, which is presented in the following.[3] We also ran a BGLMM including word concreteness as a factor, but

[3] We have also analysed the accuracy data with a standard mixed ANOVA. The pattern of significant and non-significant effects matched that of credible and non-credible effects reported for the BGLMM. Note however, that analysing accuracy data with a linear model (such as ANOVA) is highly discouraged and can lead to spurious results (e.g., Jaeger, 2008).

that factor had only a main effect, without entering into any interactions, and therefore we present the simpler model without concreteness.

A first question was whether our manipulation of processing half of a memory list had an effect on memory. The analysis for immediate memory supported an effect of our manipulation, as it showed a credible main effect of processing, implying that participants had better memory for items that were processed again after initial encoding than for items from the not-processed triplets (see Table 1 and Figure 2, upper panel). This was also true for delayed memory: Items that were processed again after encoding were better remembered than not-processed ones (see Table 2 and Figure 2, lower panel).

In what follows, we first report effects of refreshing and of elaboration on immediate memory, followed by the effects of these two processing manipulations on delayed memory. In each section, we first focus on the effects of each processing manipulation within the core design, drawing on the joint analysis of both experiments. Next, we ask how each experimental condition of the core design compares to the baseline conditions of Experiment 2. We compared the immediate memory performance in each processing condition of Experiment 2 to the baselines using pairwise comparisons of the difference distributions taken from the second BLGMM described above. **Fehler! Verweisquelle konnte nicht gefunden werden.** shows the estimated response probabilities in the immediate and delayed memory test in the six conditions of Experiment 2. The analysis revealed that there was no difference between the two baselines, indicating that free time after encoding had no effect on memory (neither immediate, $\Delta = 0.02$,

95% HDR = [-0.03, 0.07], nor delayed, Δ = 0.04, 95% HDR = [-0.01, 0.10])[4]. We therefore pooled the baselines for all following comparisons to the processing conditions reported below.[5]

8.6.3 Working Memory

Refreshing effects on working memory. We first tested how the effect of refreshing a subset of words in working memory compares to the effect of repeated reading of these words. This is the comparison through which Johnson and colleagues evaluated the effect of refreshing on delayed memory (Johnson et al., 2002; Raye et al., 2007). The BGLMM shows a credible main effect of repeat/refresh (Table 1) but in the direction of an overall advantage for trials with a repeated triplet over trials with a refreshed triplet. Thus, the direction of this effect is in opposite direction to what was observed by Johnson et al. for delayed memory. This main effect was further qualified by the two-way interaction of processing and repeat/refresh, indicating that repeated words benefited more from being processed again than refreshed words did. Nevertheless, the effect of processing appeared for both, repeated words (Δ = 0.21, 95% HDR = [0.18, 0.24]) and refreshed words (Δ = 0.10, 95% HDR = [0.07, 0.13]).

Next, we compared the effect of guided refreshing to the baseline level of memory (i.e., pooled short and long baseline), separately for processed and not processed words. Whereas the processed words in the repeat condition were remembered better than the words in the baseline conditions (Δ = 0.18, 95% HDR = [0.14, 0.21]), the processed words in the refresh condition were not – memory for refreshed words was about equal to memory in the baseline conditions (Δ

[4] For all pairwise comparisons reported in the text results are on the probability scale (i.e., Δ = 0.02 corresponds to an effect of 2%).

[5] In other words, all comparisons against the baseline reported below are based on the second BGLMM. To obtain the results, we averaged the posterior samples of the short and the long baseline and used this distribution for the calculation of the difference distributions for each pairwise comparison.

= 0.03, 95% HDR = [-0.005, 0.06]). Further, the comparison of the not-processed triplets in the refreshing condition to the baseline suggests that refreshing a subset of items in working memory harms the memory for the remaining items (i.e., the not-processed triplet; Δ = -0.08, 95% HDR = [-0.11, -0.04]). This did not appear to be the case for the not-processed triplet in the repeat condition (Δ = -0.03, 95% HDR = [-0.07, 0.00]). To summarize, whereas repeating a subset of words in working memory boosts their availability in memory above a no-processing baseline, refreshing merely maintains memory at the baseline level, while at the same time not-refreshed words drop below the baseline level.

Effects of elaboration on working memory. In the BGLMM for the core design of both experiments, there was no credible evidence for a main effect of elaboration on working memory performance, or for any of the interactions involving elaboration (see Table 1 and **Fehler! Verweisquelle konnte nicht gefunden werden.**). However, the result is not completely clear-cut. For the main effect of elaboration, 94% of the posterior mass of the effect was in the direction of better memory in the elaboration than the no-elaboration conditions. For the interaction of repeat/refresh with elaboration, 97% of the posterior indicated better memory in the repeat with elaboration than in the repeat without elaboration condition. For the three-way interaction of processing, repeat/refresh, and elaboration, 78% of the posterior indicated that the processed words in the repeat-with-elaboration condition tended to be recalled better than words in all other conditions (see Figure 2). Note that the direction of these effects was the same as the ones observed for delayed memory (compare Table 1 to Table 2) but appeared to be somewhat smaller.

Table 1 The posterior effect estimates and their 95 % HDRs of the generalized linear mixed model for binomial response variables for the immediate serial memory data of Experiments 1 and 2.

	mean parameter on probit scale	**95% HDR**
(Intercept)	0.46	[0.37, 0.56]
processing	0.32	[0.27, 0.37]
repeat/refresh	-0.20	[-0.26, -0.15]
elaboration	-0.04	[-0.09, 0.01]
processing * repeat/refresh	-0.18	[-0.23, -0.13]
processing * elaboration	-0.02	[-0.07, 0.03]
repeat/refresh * elaboration	0.04	[-0.02, 0.10]
processing * repeat/refresh * elaboration	0.04	[-0.02, 0.11]

Note. Credible effects, defined as HDRs excluding zero, are printed in bold.

Notwithstanding the absence of credible effects involving elaboration, we next compared memory for words in the elaboration conditions, as well as memory for words in the conditions without elaboration, separately against the (pooled) baseline in Experiment 2. Words in the elaboration conditions were better remembered than the baseline ($\Delta = 0.04$, 95% HDR = [0.01, 0.07]). Again, this effect was mainly driven by the words in the repeat-with-elaboration condition (see Figure 3). In contrast, the words in the conditions without elaboration showed no credible difference to the baseline ($\Delta = 0.03$, 95% HDR = [-0.01, 0.05]). In summary, the results do not provide convincing evidence for an effect of elaboration on working memory. If anything, the effect was small and confined to the condition in which the words were repeated.

8.6.4 Long-term memory

Effects of refreshing on long-term memory. The joint BGLMM of both experiments (see Table 2) revealed no evidence for a main effect of repeat/refresh on delayed memory performance, and also no evidence for an interaction of processing with repeat/refresh. Hence, contrary to the findings of Johnson and colleagues, refreshing did not lead to better long-term memory than repeated reading. We nevertheless compared words in the refresh and repeat

conditions separately against the baseline. The comparisons with the (pooled) baseline conditions indicates that the processed triplets of the refreshing without elaboration condition were remembered at about the same level as the baselines (Δ = 0.04, 95% HDR = [-0.01, 0.09]), as was memory for the processed words of the repeat without elaboration (Δ = 0.00, 95% HDR = [-0.05, 0.05]). Note that the above pattern of results also holds for a lenient score of performance in the delayed memory task, counting all responses showing correct item memory (i.e. the target, same-list items, and other-list items) as correct responses. To summarize, our results provide no evidence for an effect of refreshing on long-term memory.

Long-term memory effects from elaboration. The analysis of the delayed memory data revealed evidence for a main effect of elaboration, and an interaction of repeat/refresh with elaboration. Follow-up analyses of the interaction revealed that the elaboration effect appeared for words in the repeat conditions (*repeated with elaboration* vs. *repeat without elaboration:* Δ = 0.10, 95% HDR = [0.06, 0.14]), but not in the refresh conditions (*refresh with elaboration* vs. *refresh without elaboration*: Δ = 0.01, 95% HDR = [-0.03, 0.05]). In sum, memory was better for trials with instructed elaboration, but only when elaboration was accompanied by repetition. Furthermore, in Experiment 2 the triplets in the repeat-with-elaboration condition (Δ = 0.07, 95% HDR = [0.02, 0.11]) showed better memory than the (pooled) baseline. Likewise, the triplets in the refreshing with elaboration condition showed better memory than the baseline (Δ = 0.05, 95% HDR = [0.004, 0.08]). Together this suggests a beneficial effect of elaboration for long-term memory compared to conditions without any processing instruction, or without time for processing. None of the other differences to the (pooled) baseline were credible. The above evidence speaks for a beneficial effect of elaboration on long-term memory.

Table 2 The posterior effect estimates and their 95 % HDRs of the generalized linear mixed model for binomial response variables for the delayed memory data of Experiments 1 and 2.

	mean parameter on probit scale	**95% HDR**
(Intercept)	-0.35	[-0.42, -0.27]
processing	0.11	[0.06, 0.16]
repeat/refresh	0.04	[-0.01, 0.10]
elaboration	-0.10	[-0.16, -0.05]
processing * repeat/refresh	-0.07	[-0.15, 0.00]
processing * elaboration	-0.05	[-0.12, 0.03]
repeat/refresh * elaboration	0.12	[0.05, 0.19]
processing * repeat/refresh * elaboration	0.02	[-0.08, 0.11]

Note. Credible effects, defined as HDRs excluding zero, are printed in bold.

8.7 Discussion

Refreshing and elaboration have been proposed as control processes on the contents of working memory, potentially serving to maintain them in working memory and to lay the ground for successful long-term memory formation. Our main goal was to investigate the impact of experimentally induced refreshing and elaboration on working memory and episodic long-term memory. Comparing the effect patterns of refreshing and elaboration should help to answer the question whether the two processes are distinguishable. In the following, we will first discuss the effects of both processes on long-term memory and second their effects on working memory.

8.7.1 How do elaboration and refreshing affect long-term memory?

Long-term memory retrieval was highest for repeated items that were elaborated, confirming that participants followed the elaboration instruction, and replicating the long-term memory benefit of elaboration. Furthermore, elaboration of repeated words resulted in long-term memory benefits compared to the level of memory directly after encoding (short baseline) and

compared to the level after additional free time after encoding (long baseline), demonstrating that instructed elaboration added something over and above people's spontaneous encoding behaviour.

Refreshing showed no effects on long-term memory. We need to ask why, in contrast to the present study, previous studies reported long-term memory benefits for refreshing (Johnson et al., 2002; Johnson, Mitchell, Raye, & Greene, 2004; Raye, Johnson, Mitchell, Greene, & Johnson, 2005; Raye, Johnson, Mitchell, Reeder, & Greene, 2002). Some methodological differences might have contributed to the discrepancy between our finding and those of Johnson and colleagues (Johnson et al., 2002; Raye et al., 2002). First, Johnson and colleagues tested episodic long-term memory through a standard yes-no item-recognition test, whereas we used a 4-alternative forced-choice test tapping both item and relational memory. We do not think that this explains the different outcomes because we also found no effect of refreshing on a lenient score for delayed memory that reflected only item memory.

Second, the present study asked participants to refresh three items in working memory, whereas previous studies included refreshing memory sets of only one to two items per trial. Therefore, it is possible that refreshing benefits on long-term memory occur only when working memory load is low, so that people have a good chance to remember the words when they are asked to refresh them. Against that possibility speaks the study by Souza et al. (2015), which showed (short-term) memory refreshing benefits with six-colour arrays. Future research has to clarify whether refreshing verbal material only benefits long-term memory when the load on working memory is low.

A third discrepancy between the present study and previous research is that the present memory test was not incidental but well expected by the participants. The mere knowledge about a delayed memory test might have changed the participants' behaviour in our experiment.

Furthermore, previous research by Johnson and colleagues never included an immediate memory test, as was the case here. Thus, we might have created a testing effect for all items, whereas Johnson et al.'s method might have created a testing effect only for the refreshed item. The instruction of previous studies to “think of” included also the instruction to recall the refreshed item aloud, which confounded refreshing with recall. The increased long-term memory for refreshed items in previous studies therefore could have been due to recall – in line with the testing effect (Roediger & Karpicke, 2006) – rather than due to refreshing per se. One possible conclusion from our study is that the refreshing benefit observed by Johnson et al. can only be observed for long-term memory with their experimental design and is not universal. At least, our study demonstrates a clear boundary condition of said refreshing benefit.

8.7.2 Does elaboration benefit working memory?

As elaboration benefited long-term memory, in this as well as in previous studies, we were interested also in its immediate effects on memory. Despite its clear advantage at the delayed test, there was no compelling evidence for an effect of elaboration on immediate memory. If anything, elaboration tended to be helpful only when the to-be-remembered words were presented again in the retention interval – which is typically not the case in tests of working memory. Our findings therefore fail to provide experimental support for the conclusion from previous correlational studies, which found that higher performance on complex span tasks was related to individuals’ use of so-called "normatively effective" elaboration strategies such as imagery and sentence generation (Bailey, Dunlosky, & Kane, 2008; Bailey, Dunlosky, & Hertzog, 2009; Bailey et al., 2011; Dunlosky & Kane, 2007). Reasons for this could be due to the present study using a simple span paradigm and previous research relying on complex span tasks. Bailey et al. (2011), following Unsworth & Engle (2007), have argued that in complex span,

retrieval comes mostly from "secondary memory" (i.e., episodic LTM), and they have provided initial evidence that self-reported elaboration predicted performance only on those trials of a simple-span test that also involved retrieval from episodic LTM. Therefore, our findings could be reconciled with those of previous strategy self-report studies by assuming that elaboration improves only episodic LTM. An alternative explanation, of course, would be the reverse causality: If participants have good memory, they have more information in memory to elaborate on. They would subsequently also show better LTM.

8.7.3 How does refreshing affect working memory?

When a subset of the words in working memory is refreshed, the refreshed words are remembered better in an immediate test than the words from the not-refreshed subset. Therefore, refreshing is effective in prioritizing the refreshed representations within working memory. This effect of refreshing, however, is weaker than that of repeating the words. Our results from Experiment 2 further show that, in contrast to repeating, refreshing a subset of items in working memory did nothing to improve memory for the refreshed items relative to the baselines without any instructed processing. At the same time, the remaining items (i.e., the not-processed triplet) were remembered worse compared to the baselines. In sum, refreshing a triplet of words in working memory has no beneficial effect on immediate memory beyond maintaining performance at the level that is achieved when tested immediately after encoding. At the same time, refreshing some items negatively affects the remaining items in memory.

Our study was strongly inspired by a recent direct experimental manipulation of refreshing (Souza et al., 2015), which showed refreshing benefits for visual material in working memory. So, why did we not find a refreshing benefit above baseline level? One explanation could be that in our study items were only refreshed once. Souza et al. found that the beneficial

effect of refreshing increases with its frequency: The more refreshing opportunities an item receives, the higher the probability of recalling that item. Nevertheless, the discrepancy between our results and those of Souza et al. (2015) is not that big: The Souza et al. refreshing-frequency effect is a comparison between memory for items receiving 0, 1, and 2 refreshings within a memory set. This comparison is analogous to that between processed vs. not-processed triplets in our study, for which we obtained a credible difference: Within the refreshing condition, processed triplets (refreshed once) were recalled better than not-processed triplets. One difference still remains: In Souza et al.'s study the refreshed items were recalled *better* than the baseline conditions without any processing instruction, which was not the case in our Experiment 2. This difference between baseline and refreshing conditions, however, was not found in a more recent study of guided refreshing of visual memory items (Souza & Oberauer, 2017a). Taken together, the previous studies and the present experiments converge on the following conclusion: Refreshed items in working memory are remembered better than not-refreshed items within the same trial, but when these refreshed items are compared to baseline performance in separate trials without any refreshing instruction, refreshing yields no beneficial effect.

Apart from the Souza et al. studies discussed above, most studies on the effect of refreshing on WM performance have used complex span tasks. One could argue that the lack of a beneficial effect of refreshing in our experiments was because we did not use a complex span task. However, there is no theoretical reason why refreshing should be limited to complex span paradigms. One effect often attributed to refreshing, the McCabe effect, has recently been shown to arise from opportunities to refresh in both simple-span and complex-span tasks (Souza & Oberauer, 2017b). We therefore see no reason why our findings should not generalize to other paradigms, including complex span.

Our results are consistent with the idea that refreshing maintains representations in working memory while not-refreshed representations are forgotten over time. On this assumption, refreshing can at best be expected to keep performance at the level obtained at an immediate test (i.e., the short baseline of Experiment 2). This was the case for the refreshed triplets, whereas memory for the not-refreshed triplets dropped below that level. During a retention interval without instructed processing (i.e., the long baseline of Experiment 2), participants can be assumed to refresh all six words. One might wonder why concentrating the entire time on refreshing a subset of three words did not lead to better memory for these three words than distributing the same refreshing time over all six words. Whereas our data showed a trend in that direction (compare the refreshed triplets to the long baseline in Figure 3), there was no statistical evidence supporting it. This could be explained by assuming that refreshing is so efficient that it can maintain the memory strength of all six words, and concentrating refreshing on a subset of three items means to refresh them more than necessary, without appreciable further benefit.

Our results are also consistent with the alternative view that refreshing a subset of items prioritizes these items in working memory over the not-refreshed ones. Prioritization could mean that these items are strengthened, while nothing happens to the not-refreshed items. At test, when items compete for retrieval, the not-refreshed items tend to be blocked by their stronger competitors. Again, one might ask why selectively strengthening a subset of items does not improve memory for them above the baseline conditions. Again, the answer could be that strengthening has diminishing returns, so that the benefit of strengthening some items beyond their initial level is smaller than the cost for the remaining items.

8.7.4 Are refreshing and elaboration distinct processes?

If refreshing and elaboration are the same, then the pattern of effects of both processes on working memory and on long-term memory should be the same. The effects on working memory are not informative in this regard: If elaboration were different from refreshing, we could expect it to add something to memory performance over and above mere refreshing – it did not. This could mean that when asked to refresh a set of items, people already elaborate them, so that additionally instructing them to elaborate the items makes no difference. However, elaboration also did not add much to mere re-reading in the repeat condition either. Therefore, it could also be that refreshing and elaboration are different processes, but elaboration simply has no effect on immediate memory.

Evidence for a distinction of refreshing and elaboration lies in the long-term memory results: Compared to the mere re-reading the words in the repeat condition, elaboration benefited memory after a delay, whereas refreshing did not. Likewise, in comparison to the baseline conditions of Experiment 2, elaboration improved delayed memory whereas refreshing did not. It is particularly telling that elaboration had this beneficial effect only in conjunction with repeating, but not in conjunction with refreshing. Apparently, the need to attend to a set of words in working memory – as opposed to in the environment – undermines the effectiveness of elaboration. This is not what we would expect if refreshing and elaboration were two terms for the same process.

8.7.5 Conclusion

We replicated the beneficial effect of elaboration on long-term memory. Contrary to previous findings, we found no such effect for refreshing, suggesting that refreshing and elaboration are not the same process. Further, neither refreshing nor elaboration did much to

improve working-memory performance, compared to no processing of the memory representations after encoding. Whatever the working-memory system does spontaneously – whether engaging in some maintenance activity, or doing nothing – appears to be hard to beat by any experimentally induced process.

9. Dissociating refreshing and elaboration by their neural signatures and their effects on working memory and long-term memory in young and old adults

Lea M. Bartsch[1], Vanessa M. Loaiza[2], Lutz Jäncke[1], Klaus Oberauer[1] & Jarrod Lewis-Peacock[3]

[1] University of Zurich and University Priority Program Dynamics of Healthy Aging

[2] University of Essex

[3] University of Texas at Austin

Status: Submitted to eLIFE (September 5th, 2018)

Authors' contributions

LB reviewed the literature, designed the study, programmed the task, organized data collection, analyzed and interpreted the results and wrote the manuscript

VL reviewed the literature, conceptualized the research question, and revised the manuscript

LJ conceptualized the research question, and revised the manuscript

KO designed the research question, supervised and discussed LB's contributions, revised the manuscript

JLP supported the data analysis and interpretation, supervised and discussed LB's contributions, revised the manuscript

Author Note

We thank Lorena Hell, Danielle Pessach and Raffael Schmitt for helping with data collection. The data and the analysis scripts can be accessed on the Open Science Framework (https://osf.io/p2h8b/). The project was funded by the University Priority Program Dynamics of Healthy Aging.

8.1 Abstract

Maintenance of information in working memory (WM) is assumed to rely on *refreshing* and *elaboration*, but clear mechanistic descriptions of these cognitive processes are lacking, and it is unclear whether they are simply two labels for the same process. This fMRI study investigated the extent to which refreshing, elaboration, and repeating of items in WM are distinct neural processes with dissociable behavioral outcomes in WM and long-term memory (LTM). Multivariate pattern analyses of fMRI data revealed differentiable neural signatures for these processes, and the degree of neural separation within an individual predicted their memory performance. The benefit of refreshing items in WM increased as its neural signature became more similar to repetition. Elaboration improved LTM, but not WM, and this benefit increased as its neural signature became more distinct from repetition. This demonstrates that refreshing and elaboration are separate processes that have predictable contributions to memory performance.

8.2 Introduction

Working memory (WM) is a system for holding a limited amount of information available for processing (Baddeley, 1986), whereas episodic long-term memory (LTM) stores information permanently with presumably unlimited capacity (Tulving, 1972). WM and LTM are highly correlated constructs, and models on their relation suggest that how information is processed in WM strongly affects how well it is maintained in LTM. (D'Esposito & Postle, 2015; Eriksson, Vogel, Lansner, Bergström, & Nyberg, 2015; Lewis-Peacock & Postle, 2008; Ranganath, 2006; Ranganath & Blumenfeld, 2005; Ranganath, Cohen, & Brozinsky, 2005; Crowder, 1982; Melton, 1963; Nairne, 1990, 2002; Cowan, 1995; Oberauer, 2002). Thereby, two control processes on information in WM have been argued to contribute to encoding in episodic

LTM: *refreshing* and *elaboration*. The aim of the present study is to investigate (1) whether refreshing and elaboration are neurally and behaviorally distinguishable processes, (2) how they affect WM and episodic LTM performance, and (3) to what extent age differences in these processes are responsible for memory deficits in older adults.

8.2.1 Refreshing and elaboration: Definitions and behavioral impacts on WM and LTM

Refreshing is understood as briefly thinking of a stimulus just after it is no longer physically present but while its representation is still active (Johnson, Reeder, Raye, & Mitchell, 2002). The process was introduced as a general attention-based mechanism for enhancing and prolonging the activation of memoranda in WM (see Camos et al., 2018 for a review), thereby improving WM (Souza & Oberauer, 2016; Souza, Vergauwe, & Oberauer, 2018; but see for contradicting evidence Bartsch, Singmann, & Oberauer, 2018) and episodic LTM (Johnson et al., 2002). Researchers debate how refreshing operates, with views depending on their understanding of the WM-LTM relationship and what constitutes a WM representation. Some argue that refreshing strengthens content-context bindings and thereby creates stronger retrieval-cues for WM and LTM (Lewandowsky & Farrell, 2008; Loaiza & McCabe, 2012; Oberauer & Hein, 2012; Oberauer & Lewandowsky, 2011). Others understand refreshing as a mechanism for preserving representations against decay (Barrouillet & Camos, 2015; Lemaire, Pageot, Plancher, & Portrat, 2017).

Still another possibility is that the purported benefits of refreshing simply reflect elaboration. Elaboration refers the act of deeply processing information, especially with regard to its semantic and meaningful characteristics (Craik & Lockhart, 1972; Greene, 1987; Klatsky,

1988), and, as consequence, reliably improves episodic LTM (e.g., Craik & Tulving, 1975; Gallo, Meadow, Johnson, & Foster, 2008). The benefits of elaboration for WM are more mixed: Correlational studies show a positive relationship between elaborative strategies and verbal WM recall (Bailey, Dunlosky, & Kane, 2008, 2011; Dunlosky & Kane, 2007; Kaakinen & Hyönä, 2007) and some experimental work has shown that semantic compared to shallow processing of the memoranda yields greater WM recall (Loaiza, McCabe, Youngblood, Rose, & Myerson, 2011; Rose, Buchsbaum, & Craik, 2014; Rose, Craik, & Buchsbaum, 2014). Conversely, other work has shown unique benefits of elaboration for episodic LTM and not WM (Bartsch, Singmann, & Oberauer, 2018; Loaiza & Camos, 2016; Rose & Craik, 2012; Rose et al., 2010). Bartsch and colleagues showed that elaboration benefited LTM, but refreshing did not, and neither elaboration nor refreshing benefited WM. Furthermore, Loaiza and Camos (2018) showed that the benefit of semantic retrieval cues for verbal WM was independent of factors intended to manipulate refreshing. These studies collectively demonstrate that any refreshing benefit in WM does not result from interacting with LTM representations, thereby differentiating refreshing from elaboration. Despite these findings, it may be the case that refreshing and elaboration produce different outcomes that are derived from the same underlying process. To address this, we turn to neuroimaging data.

8.2.2 Refreshing vs. elaboration: Neural correlates

Table 1 shows an overview of all the reported regions associated with refreshing and/or elaboration. Refreshing has been associated with activity in the left dorsolateral prefrontal cortex (dlPFC, BA 8/9) and better LTM recognition compared to repeating the stimuli, and activity in the dlPFC during refreshing predicted subsequent LTM (Johnson, Raye, Mitchell, Greene, & Adam, 2003; Raye, Johnson, Mitchell, Greene, & Johnson, 2007; Raye, Johnson, Mitchell,

Reeder, & Greene, 2002) A meta-analysis (Johnson et al., 2005) identified frontal regions, specifically left dlPFC (BA 9/46), ventrolateral PFC (vlPFC, BA 44/45/47), and the left anterior PFC (BA 10) as associates of refreshing various stimulus material.

Although the dlPFC has been suggested to underlie refreshing, its activation has also been shown to predict subsequent LTM in studies of elaboration (or "relational encoding") wherein the semantic relationship between two items is elaborated upon (e.g. Blumenfeld & Ranganath, 2007). For the ease of the reader, we will refer to *relational encoding* as *elaboration* from now on. The neural correlates of elaboration have not always been that specific or limited to the dlPFC: earlier studies have more generally associated the lateral PFC with semantic elaboration (e.g., Kapur et al., 1994; Wagner et al., 1998) and relational elaboration (e.g., Addis & McAndrews, 2006; Fletcher, Shallice, & Dolan, 2000; Murray & Ranganath, 2007). Yet, numerous studies have associated the dlPFC with elaboration and subsequent memory effects (Blumenfeld, Parks, & Yonelinas, 2010; Blumenfeld & Ranganath, 2007; Davachi et al., 2001; Ragland et al., 2012). Collectively, this evidence suggests that elaboration of the memoranda in WM is what makes the dlPFC important for LTM.

Despite the neural similarities observed for refreshing and elaboration, there are some important dissimilarities in the methods used to study these processes. First, the neural correlates of refreshing have been studied for single items only, with no instructed elaboration (e.g. Johnson et al., 2005; Raye et al., 2007, 2002), and this item-specific neural processing was localized almost exclusively to left lateral dlPFC. Conversely, elaboration studies have used multiple items, such as pairs (Blumenfeld et al., 2010) or triplets of words (e.g. Blumenfeld, 2006; Davachi, Maril, & Wagner, 2001), and localized the associated activity to the bilateral dlPFC. Second, the refreshing studies have relied on incidental encoding, wherein participants

are not informed of the upcoming memory test, whereas the elaboration studies employ intentional encoding. Thus, clarifying the underlying neural processes of refreshing and elaboration requires greater consistency between the methods used to investigate them.

8.2.3 Refreshing and elaboration: Age effects

Past research has provided extensive evidence that episodic LTM declines with age (e.g., Hoyer & Verhaeghen, 2006; Naveh-Benjamin & Old, 2008; Zacks, Hasher, & Li, 2000), but the source of the deficit is still under debate. One view is that WM maintenance processes and recruitment of corresponding brain areas decline in older age (Hoareau, Lemaire, Portrat, & Plancher, 2016; Plancher, Boyer, Lemaire, & Portrat, 2017; Smith, 1980). For instance, it has been shown that older adults exhibit reduced refreshing-related brain activity in the left dlPFC and reduced refreshing benefits for episodic LTM relative to young adults (Johnson, Mitchell, Raye, & Greene, 2004; Raye, Mitchell, Reeder, Greene, & Johnson, 2008). Another possibility is that older adults are less likely than younger adults to engage in elaboration, thereby resulting in deficient retention (Smith, 1980). For example, some work has shown older adults are able to capitalize on experiment-administered elaborative strategies but show deficiencies in generating elaborative strategies themselves (Rankin & Collins, 1985, see also Kamp & Zimmer, 2015). A meta-analysis reported that age-related differences in subsequent memory are associated with under-recruitment of the occipital and fusiform cortex as well as an over-recruitment of medial and lateral regions of PFC and parietal lobe (Maillet & Rajah, 2014). These findings suggest inefficient recruitment of brain regions that are important for elaboration, thereby leading to age-related memory deficits.

8.2.4 The present study

The goal of the present study was to investigate to what extent elaboration and refreshing are separable processes, given their neural overlap as well as their similar proposed beneficial effects for memory. So far, only one study has investigated both processes in one experiment, and the behavioral results demonstrated that the processes have divergent contributions to LTM (Bartsch et al., 2018). We aimed at extending this previous study by not only investigating whether refreshing and elaboration are distinct in their contribution to WM and LTM formation, but also whether they are supported by separable neural activation patterns. Furthermore, we aimed to investigate their impact on age-related memory deficits.

We applied multivariate pattern analyses (MVPA; e.g., Haxby, Connolly, & Guntupalli, 2014; Haxby et al., 2001; Haynes & Rees, 2006; Lewis-Peacock & Norman, 2014; Norman, Polyn, Detre, & Haxby, 2006) to fMRI data of young adults and older adults performing the word list encoding task of Bartsch et al. (2018). This analysis approach allowed us to evaluate whether brain activity patterns associated with refreshing items and with elaborating items in WM could be differentiated. These neural measures were then linked to behavioral outcomes on tests of both WM and LTM. MVPA is especially sensitive to detecting fine-grain differences between neural activation patterns that are not detectable using conventional analyses (N Kriegeskorte & Bandettini, 2007).

If refreshing and elaboration are two labels for the same process, then the pattern of behavioral effects should be similar for WM and on LTM, and the patterns of brain activity supporting these processes should be indistinguishable. If refreshing and elaboration are distinct processes, they should have different behavioral effects and separable patterns of neural activation. As elaboration has previously been shown to benefit LTM compared to re-reading the

memory material, we expected the greatest LTM benefit when the neural signatures of elaboration and repetition were most distinct. On the other hand, because re-reading has been shown to benefit WM *more* than refreshing, we expected the greatest benefit of refreshing when these processes were most similar. Finally, if older adults' memory deficits are caused by deficient uses of refreshing and elaboration, we expected their neural signatures to be less distinct and more weakly correlated with behavioral outcomes compared to young adults.

8.3 Method

8.3.1 Subjects and general procedure

We recruited 30 healthy, right-handed young adults (15 females; mean age = 24.2, SD = 2.97 years) from the student population of the University of Zurich as well as 27 healthy, right-handed older adults from the community (13 females; mean age = 69, SD = 3.47 years). Handedness was measured through observation of the writing hand. Subjects were screened for their ability to undergo a magnetic resonance imaging session. Furthermore, they completed the Digit–Symbol Substitution test (DSS; Wechsler, 1982), serving as an indicator of processing speed, and the mini-mental-status examination (MMSE; Folstein, Folstein, & McHugh, 1975) to screen for cognitive impairment. All subjects performed a WM task while being scanned with a 3-T MRI scanner, and subsequently an LTM task outside the scanner. The session ended with a computerized version of a vocabulary test (Lehrl, 2005), a marker test for crystalized intelligence. The study was approved by the ethical review board of the canton of Zurich. The participants were compensated with either 60 Swiss Francs (about 60 USD) or partial course credit for the two-hour session.

8.3.2 Paradigm

The paradigm is the same as reported in a recent study (Bartsch et al., 2018), adapted for use in the MRI scanner. We asked participants to remember six nouns in serial order (see

Figure 2). After list presentation, either the first three words or the last three words were to be processed again in one of four ways, depending on the experimental condition. During encoding it was not predictable which half of the items would have to be processed. In the *repeat* condition, the three words appeared again sequentially on the screen, and the subjects had to simply re-read them silently (we will also refer to this as the *re-read* condition). In the *refreshing* condition, the to-be-processed words were replaced by refreshing prompts appearing at the same location. The subjects were instructed to "think of" the corresponding words as soon as the prompts were shown. In the *elaboration* condition, the three to-be-processed words were shown again sequentially on the screen, and subjects were instructed to generate a vivid mental image of the three objects interacting. The stimuli appearing on the screen in that condition did not differ from the repeat condition, leaving the instruction to form a vivid mental image as the only difference between these conditions. Finally, in the combined *refresh-with-elaboration* condition the participants had to "think of" the words replaced by the prompts, and additionally form a vivid mental image of those items. Again, the event sequence of this condition does not differ from the refreshing condition apart from the instruction to form a mental image. Memory was tested with a four-alternatives forced-choice task, which we describe in detail below (see Section *Procedure: Working memory task*).

The experiment used a 2 x 2 x 2 x 2 (repeat/refresh [repeat, refresh] x elaboration [with elaboration, without elaboration] x processing [processed triplet, unprocessed triplet] x age [young adults, older adults]) within-subject, between-age group design. Orienting the processing

task to only a subset of the words in memory allows us to draw inferences about the effect of each of the processing conditions on memory by comparing the memory performance of the triplet of words that was not further processed after encoding to the triplet that was processed according to one of the four experimental conditions.

8.3.3 Materials

The stimuli were nouns randomly drawn from a pool of 863 German abstract and concrete nouns for each subject. The nouns were between two and 15 letters long and had a mean normalized lemma frequency of 30.81/million (drawn from the dlexdb.de lexical database).

8.3.4 Procedure: Working memory task

The sequence of an experimental trial is illustrated in

Figure 2. The experiment was performed using Presentation® software (Version 18.0, Neurobehavioral Systems, Inc., Berkeley, CA, www.neurobs.com). The six to-be-remembered words in each trial were sequentially presented in boxes from top to bottom on the screen, each for 500 ms. Depending on the experimental condition, a cue was presented 1000 ms after the last memory item, indicating whether the first half or the second half of the list had to be processed again. In the elaboration and repeat conditions, each word in the to-be-processed triplet was shown again for 1400 ms, followed by a 600 ms inter-stimulus interval. In the refreshing and refreshing-with-elaboration conditions, each to-be-processed word of a triplet was replaced by a refreshing prompt (*#?#?#*) in its corresponding box, and participants were instructed to "think of" the word in that box. In the elaboration and refresh-with-elaboration conditions, participants

were additionally instructed to form a vivid mental image of the three words interacting with each other.[6]

After processing the words in the cued triplet, participants' memory for each list item was tested in their order of presentation using a 4-alternative forced-choice procedure. For each tested item, four words were presented from which the subject could choose the correct word in the currently tested list position with a button press. All test sets included the following four response options: the target (i.e., correct) word, one lure from the same triplet of words within the present list, one lure from the other triplet of the present list, and one new word. This choice had to be made for each of the serial positions successively and with a time limit of 2500 ms for the young and 3500 ms for the older adults per serial position to ensure controlled timing for the fMRI image acquisition. We applied this 4-alternatives forced-choice recognition task in order to test both memory for items (i.e., discriminating between items that have been presented in the current memory list and new items) and for serial order (i.e., discriminating between the item in the tested position and other list items).

Within each block of four trials, the same type of processing was instructed throughout, and a screen repeating the instructions of the particular condition was shown prior to the beginning of each block. The order of the condition blocks was randomized between subjects. Each of the four fMRI runs consisted of four blocks, one for each condition (with 4 trials per block as described above).

[6] The timing parameters where chosen based on a pilot experiment with young adults, which allowed participants to process the items in each of the 4 experimental conditions in a self-paced mode. The mean processing times (PT) where PT = 1419 ms in the repeat without elaboration condition, PT = 1491 ms in the repeat with elaboration condition, PT = 1197 ms in the refreshing without elaboration, and PT = 1198 ms in the refreshing with elaboration condition.

8.3.5 Procedure: Long-term memory task

After leaving the scanner participants were brought into a separate room, where they performed the computerized LTM task. We assessed participants' LTM for the words they had encoded for the WM tests throughout the experiment. To this end, we presented in each trial the first word of a triplet from one of the studied memory lists. We asked participants to choose, from four different options, the word that had followed the target word in that triplet. The probe words included the correct word (i.e., which could be either the word in the second or third position of the target triplet for the first prompt, and the fifth or sixth word for the second prompt), two words from another list, and a new word. This allowed us to keep the format of the LTM test very similar to the WM test, and furthermore to compare in each trial the memory performance for words from the processed and from the unprocessed triplets. As in the WM test, the LTM test also provided information about both item memory (i.e., which words have been presented in the experiment) and relational memory (i.e., which words have been together in a triplet). The participants were made aware of the LTM test before the start of the experiment.

8.3.6 fMRI Data Acquisition and Preprocessing

Whole brain images were acquired with the 3 T Philips Ingenia MRI scanner with a 32-channel head coil, located at the University Hospital Zurich, Switzerland. High-resolution T1-weighted images were acquired for all subjects with a Turbo field echo (TFE) sequence (8ms time repetition (TR), 3.7ms time echo (TE), 8° flip angle, 160 sagittal slices, 240 × 240 inplane, 1.0mm isotropic). Blood oxygen level-dependent (BOLD)-sensitive functional MRI data were acquired using a gradient-echo, echo planar sequence (2 s TR, 35ms TE) within a 72 × 70 matrix (32 transverse slices, 3 mm isotropic).

Following the acquisition of the structural images, four MRI acquisition runs were collected for each subject, in which they performed a 10-min block of a six-item WM task with a processing delay. fMRI data preprocessing (slice-time correction and realignment) was performed with SPM12 (Penny, Friston, Ashburner, Kiebel, & Nichols, 2011). Subjects' functional scans were aligned by realigning the first volume in each run to the first volume of the first run, and then registering each image in each run to the first volume of that run. The middle functional slice served as a reference for slice-time correction. Further, the functional volumes were co-registered to the T1 anatomical image.

8.3.7 Analysis of Behavioral Data

All data and analysis scripts can be assessed on the Open Science Framework (osf.io/p2h8b/).We analyzed the behavioral data using a Bayesian generalized linear mixed model (BGLMM) implemented in the R package *rstanarm* (Stan Development Team, 2018) following the exact analysis pipeline reported by Bartsch and colleagues (2018). The dependent variable was the number of correct and incorrect responses in each cell of the design per participant. Correct responses were defined as choosing the target item from the four alternatives. Bayesian procedures provide posterior probability distributions of the model parameters (i.e., the regression weights) that express uncertainty about the estimated parameters. The highest density regions (HDRs) of these posteriors can be used for statistical inference. A 95% HDR represents the range in which the true value of a parameter lies with probability 0.95, given model and data (Morey, Hoekstra, Rouder, Lee, & Wagenmakers, 2016). If zero lies outside the Bayesian HDR there is strong evidence for the existence of the corresponding effect. Although the strength of evidence varies continuously, for simplicity we will describe effects as "credible" if their HDRs

exclude zero. We used an MCMC algorithm (implemented in Stan; Carpenter et al., 2017) that estimated the posteriors by sampling parameter values proportional to the product of prior and likelihood. These samples are generated through 4 independent Markov chains, with 1000 warmup samples each, followed by 1000 samples drawn from the posterior distribution which were retained for analysis. Following Gelman and colleagues (2013), we confirmed that the 4 chains converged to the same posterior distribution by verifying that the $\hat{R}$ statistic – reflecting the ratio of between-chain variance to within-chain variance – was < 1.01 for all parameters, and we visually inspected the chains for convergence.

8.3.8 Generation of ROIs

We included all ROIs that were previously reported in fMRI studies investigating either refreshing or elaboration and that had shown subsequent memory effects and/or significant activation differences between repeating and refreshing or elaboration in univariate analyses (see Table 3 for details). This search was performed using the neurosynth.org database and keyword-based search in pubmed.gov. Anatomical ROIs were generated using an automated parcellation method from *FreeSurfer*. Briefly, a surface mesh model was reconstructed for each subject. Each subject's surface was then auto-parcellated based on the folding pattern of the gyri and sulci. We generated ROIs corresponding to frontal, parietal, fusiform, parahippocampal and temporal regions in this manner. We constructed combined masks using *fslmaths* to create the frontal mask, encompassing Brodmann areas 44, 45, and 47 for the inferior frontal mask, and including Brodmann areas 4, 6, 8, 9, 10 and 46 for the superior frontal mask. The parietal mask included Brodmann area 3, 7 and 40. The temporal mask consisted of the inferior-, middle-, and superior

temporal labels of the *aparc* atlas and the fusiform mask consisted of the fusiform label of the atlas. We further combined masks with *fslmaths* in various combinations of the above (e.g. frontal-parietal mask, frontal-temporal-parietal mask, frontal-fusiform mask, frontal-fusiform--parietal mask).

8.3.9 Multivariate Pattern Analyses of fMRI Data

MVPA provides greater inferential power than classical univariate approaches due to its higher sensitivity at detecting information in neural signals. As a result, MVPA has led to the successful within-category decoding of the contents of WM at an item level (LaRocque, Riggall, Emrich, & Postle, 2017) as well as the characterization of neural representations in different states of WM (Christophel, Iamshchinina, Yan, Allefeld, & Haynes, 2018; Lewis-Peacock, Drysdale, Oberauer, & Postle, 2012; Lewis-Peacock, Drysdale, & Postle, 2015; Rose et al., 2016). The sensitivity of MVPA was further established by a study demonstrating that allegedly category-selective brain regions detected in univariate analyses of the BOLD signal during the delay period of a WM task still carried patterns of activity associated with another category of information, that was currently relevant for behavior (Lewis-Peacock & Postle, 2012).

MVPA was performed in MATLAB using the Princeton MVPA toolbox (http://code.google.com/p/princeton-mvpa-toolbox). The classification algorithm used for this analysis was a L2-regularized binary logistic regression (1 vs. the others), that uses Carl Rasmussen's conjugate gradient minimization algorithm, with a penalty term of 50. The classification was performed in the anatomically defined ROIs defined above. All neural data were high-pass filtered with a cut-off of 128 seconds and z-scored across trials, within runs, before running MVPA. We performed ANOVA-based feature selection of all active voxels

within the respective ROI mask and chose the voxels that individually were able to discriminate between the three conditions (repeat, refreshing, elaboration) significantly ($p < .05$) over the course of the experiment. To avoid circularity in the data analysis, feature selection was performed separately for each iteration of the cross-validation classifier training algorithm, using only the training and testing sets independentdata in each iteration (Kriegeskorte, Simmons, Bellgowan, & Baker, 2009). This univariate feature selection technique has been shown to reliably improve classification accuracy in MVPA of fMRI (Lewis-Peacock & Norman, 2014). The pattern of activity across these feature-selected voxels was used as the input to the pattern classifier.

MVPA — Three-way problem: repeat vs. refreshing vs. elaboration.

The classification procedure used k-fold cross-validation on the data from the WM task. Preprocessed fMRI data from each 6-s processing period (three volumes) from the WM trials, after accounting for a 6-s hemodynamic lag, were used for the analysis. Our analysis scheme incorporated each functional volume (acquired over a 2-s TR) as a separate training event, so that every trial resulted in three events. Each event was associated with an array of features corresponding to BOLD signal in voxels in the ROI being used.

The k-fold cross-validation scheme (k = 4, for each of the runs) trained a classifier, separately for each participant, on the data of the three conditions (repeat, refresh and elaboration) from three runs and then used this classifier to test the data from the withheld run. This process was repeated until every run had been held out for testing. The statistical significance of classifier accuracy was evaluated by performing permutation tests on relabeled training data, in each cross-validation fold, and comparing the resulting distribution of classifier

accuracies to the true (unshuffled labels) classifier accuracy with a one-sample, one-tailed t-test. This analysis scheme was performed for every ROI. Finally, classification performance was also assessed using receiver operating characteristic (ROC) curves, which rank the classification outputs according to their probability estimates (from strongly favoring Class A to strongly favoring Class B or C) and chart the relationship between the classifier's *true positive rate* (probability of correctly labeling examples of Class A as Class A) and *false positive rate* (probability of incorrectly labeling examples of Class B or C as Class A) across a range of decision boundaries. The area under the curve (AUC) indexes the mean accuracy with which a randomly chosen pair of Class A and Class B (or C) trials could be assigned to their correct classes (0.5 = random performance; 1.0 = perfect performance).

MVPA — Two-way problems: repeat vs. refreshing and repeat vs. elaboration.

In order to assess how the neural classification of the refreshing process as well as the neural classification of the elaboration process relates to an individual's task performance, we used the repeat condition as a reference. First, we extracted classification scores from repeat and refresh trials only, using classifiers that were trained on all three processes. In order to assess the neural separability of the perceptually identical conditions of repeat and elaboration, we chose to retrieve the evidence values of the two-way problem from the three-way trained classifier the same way as we did for the repeat vs. refreshing comparison, with the difference that we did this for each individuals' mask with the *highest* classifier accuracy. The reason for this is that elaboration instruction (*"form a vivid mental image of the three words interacting with each other"*) can be implemented very differently by each subject, resulting in unique subject-specific neural signatures of elaboration. Here, our primary interest was to discover *whether* elaboration

could be identified and distinguished, rather than to identify precisely *where* in the brain the processes were implemented. Once again, we assessed classifier performance for each binary classification problem using AUC.

MVPA — refreshing with elaboration.

We were further interested in whether the patterns of activity in the brain during the combined refreshing with elaboration condition would be similar to refreshing or to elaboration or would elicit neural patterns of a completely new process.

Equivalent to the above analysis scheme we trained classifiers on all four conditions (repeat, refreshing, elaboration, and simultaneous refreshing and elaboration) and then we extracted classification scores from refreshing with elaboration trials only, using classifiers that were trained on all four processes, again in a k-fold cross-validation scheme. In order to assess the neural separability of the conditions of refreshing, elaboration, and refreshing with elaboration, we chose to retrieve the evidence values of *refresh*, *elaborate* and *refresh with elaboration* from the four-way trained classified. With this approach of training on a four-way-problem and testing on three processes, we computed the evidence values for refreshing, elaboration or the refreshing with elaboration condition. This allowed us to explore whether the brain activation patterns of the combined refreshing with elaboration events were more similar to refreshing or to elaborating or elicited a completely different process (note that classifier evidence can be construed as an estimate of the similarity between patterns of activity, with high evidence values implying low similarity of the target condition to all other conditions, and vice versa). Using k-fold cross-validation, we trained classifiers on three runs of data and tested it on the withheld run, but only on trials from the refreshing with elaboration condition. Note that it

was important to include the repeat condition in the training set to help the classifier distinguish unique process-level activity associated with the perceptually identical elaboration condition.

8.3.10 Researcher Degrees of Freedom

Analyses of neural data involve many decisions, and when these decisions are informed by the data to be analyzed, there is a risk that they are biased in favor of a desired outcome (Simmons, Nelson, & Simonsohn, 2011). Some aspects of our analysis plan (in particular, the decision to use anatomically defined ROIs for the MVPA analyses) were informed by the data of the young adults. Our analysis of the old adults' data, however, used the exact same analysis pipeline as that for the young adults without any adjustment informed by the old adults' data. Therefore, any convergent finding in both age groups can be thought of as having been directly replicated in a different population. For any finding that differs between age groups, there remains an ambiguity as to whether the divergence reflects a failure to replicate the finding in the young-adult sample, or a genuine age difference. Resolving this ambiguity requires a replication of the entire study with the present analysis plan.

8.4 Results

8.4.1 Behavioral Results

We replicated all effects of the young adults reported in a previous study (Bartsch et al. 2018). Figure 3 shows the estimated proportion of correct responses and their corresponding 95% highest posterior density regions for the immediate and delayed memory data. The posterior effect estimates are presented in Table 1 and Table 5. A first question was whether our manipulation of processing half of a memory list had an effect on memory. The credible main effect of processing on immediate and delayed memory supported an effect of our manipulation:

Participants had better memory for items that were processed again after initial encoding than for items from the unprocessed triplets (see Table 1 & Table 5 and Figure 3). There was also a main effect of age, such that older adults showed worse memory performance on tests of both WM and LTM.

Working memory performance. We first tested how the effect of refreshing a subset of words in WM compares to the effect of repeated reading of these words. This is the comparison through which Johnson and colleagues evaluated the effect of refreshing on delayed memory (Johnson et al., 2002; Raye et al., 2007). There was a main effect of repeat/refresh (Table 1), but with an advantage of repeating over refreshing. This main effect was further qualified by the two-way interaction of processing and repeat/refresh, indicating that repeated words benefited more from being processed again than refreshed words did. Nevertheless, the effect of processing appeared for both repeated words ($\Delta = 0.34$, 95% HDR = [0.31, 0.37]) and refreshed words ($\Delta = 0.12$, 95% HDR = [0.10, 0.15]). Furthermore, the factor of repeat/refresh interacted with age, indicating that older adults had a greater advantage of repeat over refreshed trials than young adults. Nevertheless, the repeat-refresh difference appeared for both, young ($\Delta = 0.16$, 95% HDR = [0.13, 0.18]) and older adults ($\Delta = 0.09$, 95% HDR = [0.06, 0.12]).

The BGLMM revealed no credible evidence for a main effect of elaboration on WM performance, or for any of the interactions involving elaboration (see Table 1).

Long-term memory performance. The BGLMM revealed evidence for a main effect of repeat/refresh on LTM performance, but as with WM, there was an advantage for repeating over refreshing (see *Table 5*). There was no evidence for any further interaction including the repeat/refresh factor. Hence, contrary to the findings of Johnson and colleagues, refreshing did not lead to better LTM than repeated reading. Note that the above pattern of results also holds for

a lenient score of performance in the LTM task, counting all responses showing correct item memory (i.e. the target, same-list items, and other-list items) as correct responses.

Furthermore, the analysis of the LTM data revealed evidence for an interaction of elaboration with age (see *Table 5*). Follow-up analyses of the interaction revealed that a beneficial effect of elaboration appeared only for young ($\Delta = 0.05$, 95% HDR = [0.02, 0.06]), but not older adults ($\Delta = -0.01$, 95% HDR = [-0.04, 0.3]). In sum, memory was better for trials with instructed elaboration than for those without, but only for the young and not the older adults. The above evidence speaks for an age-dependent beneficial effect of elaboration on LTM that is lost in older age.

To summarize, our results provide no evidence for an effect of refreshing on LTM for either age group; instead we replicated the benefit of elaboration on LTM but only for young adults.

8.4.2 MVPA Results

Young adults

Repeat vs. Refresh vs. Elaborate.

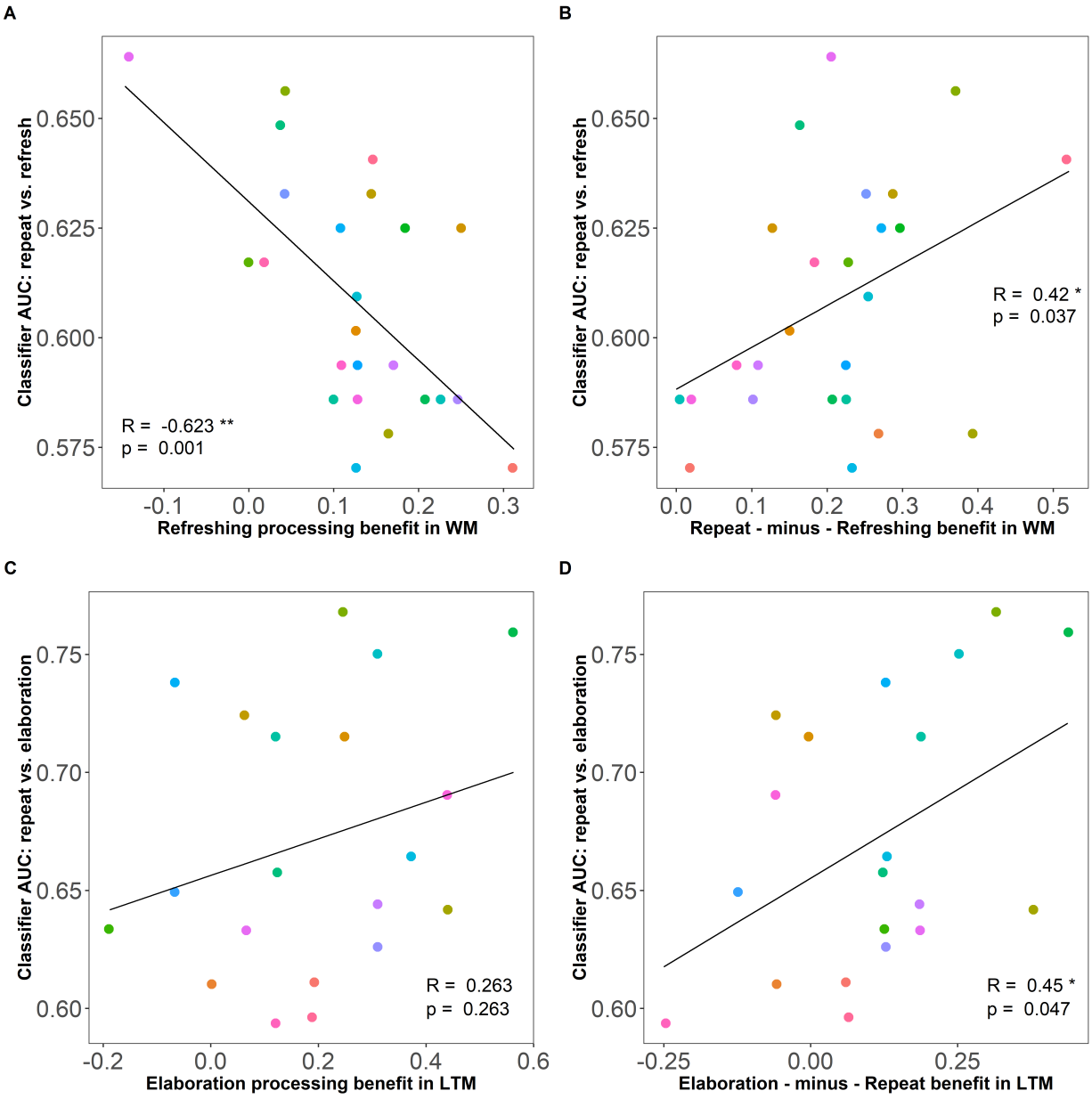
The classification scores for each individual were converted to a sensitivity score, accounting for both hits and false alarms, by computing the area under the ROC curve (AUC) for the three-way classification. For 25 of the 30 subjects classification of repeat, refresh, and elaborate processes was successful (M_{AUC} = 57.17 %, SD_{AUC} = 4.16%, i.e., significantly better than chance with $p < .05$) in at least one of the predefined anatomical ROIs (see Methods) (Figure 4b). Mean classification sensitivity for these 25 subjects in their significant ROIs was M_{AUC} = 57.17 % (SD_{AUC} = 4.16%). Data from the five remaining subjects were excluded from the subsequent analyses. A majority of the subjects (N = 17) showed good classification in the

combined bilateral frontal–fusiform–parietal regions, and the remaining subjects (N = 12) showed good classification in various subsets of this larger mask (see Figure 4). Notably, whole-brain classification was less successful than this ROI-based approached (M_{AUC} = 53.49%).

Linking neural classification to memory performance

Refresh vs. Repeat

As in the three-way classification analysis, the classification scores for each individual were converted to a sensitivity score, accounting for both hits and false alarms, by computing the area under the ROC curve (AUC) for the *refresh* vs. *repeat* classification. The classifier AUC for separating re-reading from refreshing was well above chance in the combined mask of frontal–fusiform–parietal regions and its various subsets (M_{AUC}= 61.14%, *SD* = 2.9%, chance = 50*%, p < .001-.05*; Figure 5b). These neural separability scores were then correlated with three individual behavioral metrics: (1) the *repeat processing benefit* contrasts memory for the processed triplets (i.e., the words that were re-read) vs. the unprocessed triplets in that condition, (2) the *refresh processing benefit* contrasts memory for the processed triplet (i.e., the words that were refreshed) vs. the unprocessed triplets in that condition, and (3) the *repeat-minus-refresh benefit* contrasts the memory of the processed triplets in both conditions. The *repeat processing benefit* was not significantly correlated with neural separability (r = 0.146, p = .485). A higher *refresh processing benefit* was significantly correlated with *lower* neural separability between *refresh* and *repeat* (r = -0.591, p = .0263; Figure 7



a). The more an individual benefited from refreshing parts of a list, the more neurally similar this process was to re-reading parts of a list. There was a significant positive correlation between classifier AUC scores for the *repeat vs. refresh* processes and the behavioral *repeat-minus-refresh benefit* (r = 0.419, p = .047; Figure 7b). That is, the more neurally separable *repeat* was from *refresh*, the larger the relative advantage of repeating over refreshing the words. Conversely, memory for refreshed words approximated the (higher) accuracy of repeated words when these neural processes were more similar. In summary, these results indicate that although repeating

items benefited WM performance more than refreshing did, this relative advantage was reduced in individuals for which the neural processes of repeating and refreshing were more similar. That is, refreshing benefited WM more when it appeared, in the brain, to be more like the subject was re-reading the words.

Elaborate vs. Repeat

To assess how the neural classification of the elaboration process relates to an individual's task performance, we again used the repeat condition as a reference (see Methods). For 20 of the 25 subjects who showed significant classification in the three-way analysis, the processes of re-repeating and elaboration were classifiable significantly above chance in at least one mask, with a mean *AUC* of 67.11 % (*SD* = 5.67%). The individuals' best masks formed two clusters: one of bigger masks, including frontal and temporal regions (6 subjects) and the other of just frontal regions (11 subjects). The remaining 4 subjects showed highest classifiability of *repeat vs. elaborate* in temporal regions (see Figure 6b).

Elaboration had no behavioral effect on WM, but instead showed a benefit for LTM. Therefore, our analysis focused on three behavioral contrasts in the LTM accuracy data: (1) the *repeat processing benefit* contrasts memory for the processed triplets (i.e., the words that were re-read) vs. the unprocessed triplets in that condition, (2) the *elaboration processing benefit* contrasts memory for the processed triplet (i.e., the words that were elaborated) vs. the unprocessed triplets in that condition, and (3) the *elaborate-minus-repeat benefit* contrasts the memory of the processed triplets in both conditions.

Across subjects, the relative LTM performance benefit of *elaborate–minus-repeat* was positively correlated with the classifier AUC scores reflecting the neural separation of these two

processes (r = 0.45, p < .05; Figure 7d). The more these processes were separable in the brain, the more that elaboration provided a benefit for LTM beyond simply re-reading the words. Neither the *repeat processing benefit* nor the *elaboration processing benefit* were significantly correlated with neural separability (r = 0.17, p = 0.487 and r = 0.26, p = 0.263, respectively). Although the neural separability of *elaborate vs. repeat* related to the LTM performance benefit of *elaborate-minus-repeat*, we did not find such an effect on WM (r = 0.17, p = 0.484). Similar to LTM, neither the *repeat processing benefit* nor the *elaboration processing benefit* were significantly correlated with neural separability (r = -0.12, p = 0.61 and r = -.07, p = 0.76, respectively).

Refreshing with Elaboration

The four-way classification (repeat, refreshing, elaboration, refreshing with elaboration) yielded above chance classification in 27 out of 30 subjects, with a mean accuracy of 32.68% (SD = 2.95), in the combined mask of frontal–fusiform–parietal regions or its subset mask (frontal N = 3, fusiform left hemisphere N = 2, inferior frontal N = 1, inferior frontal right hemisphere N = 2, superior frontal left hemisphere N = 1). In order to assess the neural similarity of refreshing with elaboration to the individual processes of refreshing and elaboration, we used cross-validation to train a classifier and extract classifier evidence values for the *refreshing*, *elaboration,* and the *refreshing with elaboration* conditions. The classifier was applied to the held-out testing run of each fold of the cross-validation, and classifier results were extracted for each trial of the *refreshing with elaboration* condition. The evidence values across all of these trials were highly similar for each of the three trained categories. To probe further, we calculated the proportion of trials in this *refreshing with elaboration* condition that were classified as refreshing, elaboration, or the combined condition. Over all subjects, each of the three processes

was predicted equally often. In other words, the classifier was unable to reliably distinguish the combined *refreshing with elaboration* condition from either the *refreshing* condition or the *elaboration* condition.

Older Adults

Repeat vs. Refresh vs. Elaborate.

The same analysis pipeline for the three-way problem in the young adults was subsequently applied to the independent sample of 27 older adults. For 22 subjects the classification of *repeat vs. refresh vs. elaborate* was significantly above chance in the frontal-parietal-fusiform mask or its subset masks, with a mean classifier *AUC* of 54.02 % (*SD* = 3.79%, see Figure 8b). This result indicates that the three processes were neurally separable also in older adults. Data from the five subjects for which the cross-validation classification accuracy was not significantly above chance were excluded from the subsequent analyses.

Linking neural classification to memory performance

Repeat vs. Refreshing

In order to assess how the neural classification of the processes of refreshing repeating relates to memory performance for the older adults, we again applied the same analysis approach as performed on the young adults. We retrieved the classifier *AUC* scores for separating repeating from refreshing in the 22 subjects (M_{AUC} = 61.58%, *SD* = 3.36%, $p < .001$ - .05) in the combined mask of frontal–fusiform–parietal regions or in one of its subsets (Figure 9b). The AUC scores was significantly correlated to subjects' behavioral *repeat-minus-refresh benefit*: The bigger the repeating and refreshing was (with higher performance in the repeating condition), the more separable were the two processes in the brain, represented by higher *AUC*s (r = 0.43, p = .031;

see Figure 11a). This is the same relation as we had observed in the young sample. In contrast to the young adults, we did not find a significant correlation of neural separability to the *refresh processing benefit* (r = -0.001, p = .998). Instead these neural measures were correlated with subjects' *repeat processing benefits*: Individuals with higher neural separability of re-repeating vs. refreshing showed a larger benefit for the portions of the word lists that were *repeated* (r = 0.48, p = .016, see Figure 11b).

In summary, the older adults' *repeat processing benefit* in WM performance was predictive of the neural separability of the two processes of repeating and refreshing. Diverging from young adults, the older adults' *refreshing processing benefit* was unrelated to the neural classifiability of these processes. This result might indicate a different, and less effective, strategy of refreshing was used by older adults compared with that of the young adults.

Repeat vs. Elaborate

In order to assess the neural separability of the perceptually identical conditions of repeating and elaboration in older adults, we retrieved the classifier evidence values of the *repeat* and the *elaboration* conditions from the three-way trained classifier, in this case using each individuals' mask with the highest classifier *AUC*. For 20 out of the 22 subjects that showed significant classification in the three-way analysis, the processes of repeating and elaboration were reliably separated, with a mean classifier *AUC* of 62.77 % (*SD* = 3.81, see Figure 10b). The individuals' best masks formed four clusters: one of bigger masks, including frontal and temporal regions (5 subjects) another of just frontal regions (5 subjects), one of temporal regions (7), and one of parietal regions (3).

Just as in the young adults, elaboration had no behavioral effect on WM, but in contrast to the young adults, the older adults also showed no benefit of elaboration on LTM. An individual's classifier *AUC* score was unrelated to WM performance, measured either by the *elaboration-minus-repeat benefit* (r = -0.13, p = 0.595), or by the *elaboration processing benefit* (r = 0.34, p = .146). There was also no significant correlation to LTM performance, indicated by the same contrasts (*elaboration-minus-repeat benefit*: r = -0.27, p = .244 and *elaboration processing benefit*): r = -0.15, p = .53).

8.5 Discussion

The goal of the present study was to investigate to what extent elaboration and refreshing are separable processes, given prior reports of their neural overlap as well as their similar proposed roles for WM and LTM. We aimed at investigating whether refreshing and elaboration are distinct in their contribution to WM and LTM formation, whether they elicit separable neural activation patterns in fMRI, and how they relate to age-related memory deficits. We compared the neural and behavioral results of these processes to a control condition of re-reading (repeating) the words during the delay-period of a WM task. In the following, we discuss the effects of refreshing and elaboration on WM and LTM, as compared to repeating, and we argue that these processes are distinct and have distinct consequences on memory performance in young and old adults.

8.5.1 How does refreshing affect WM and LTM?

We replicated the behavioral findings from Bartsch et al. (2018) that repeating items benefited WM performance more than refreshing did. This relative advantage was reduced when

the neural processes of reading and refreshing were more similar. That is, refreshing benefited WM the most when it appeared, in the brain, to be most like the process of re-reading the memory items. This finding contributes to the current debate of *which* aspect of a word's representation is being reactivated by refreshing: Our results indicate that it is beneficial to refresh the visual perceptual trace of the word or the lexical representation that is elicited also by reading, with these characteristics forming a part of its multi-faceted representation in WM (D'Esposito & Postle, 2015; Eriksson et al., 2015; Lewis-Peacock & Postle, 2008; Ranganath, 2006; Ranganath & Blumenfeld, 2005; Ranganath et al., 2005).

The assumption that successful refreshing primarily re-activates the visual aspects of word representations converges with recent behavioral findings suggesting that thinking of an item in a perceptual/visual manner could result in a larger refreshing benefit: Souza, Vergauwe, and Oberauer et al. (2018) presented cues to refresh memory items during the retention intervals of various WM tasks to instruct people to refresh the cued items. The frequency of refreshing each item increased memory performance for both verbal and visual-spatial stimuli, but the benefits of refreshing were larger for visual–spatial than verbal materials. Hence, although refreshing is conceptualized as a domain-general mechanism, it might be optimized for the visuospatial system.

Another explanation for the reduced neural discriminability of re-reading vs. refreshing for those individuals with larger *refresh processing benefits* could be that the memory benefit arises from another process altogether, one that may be performed similarly during re-reading and refreshing: the strengthening of retrieval cues to the target information (Lewandowsky & Farrell, 2008; Oberauer & Hein, 2012; Oberauer & Lewandowsky, 2011). Accordingly, successful refreshers could be strengthening the content-context bindings in the refreshing

condition in a similar way as on repeat trials in which they can explicitly re-encode the information. The current results cannot distinguish between these two interpretations. Nevertheless, what our results demonstrate is that successful refreshers perform this operation more similarly, and with similar behavioral advantages, as when they are afforded an opportunity to re-read the physical stimuli prior to the memory test.

8.5.2 How does elaboration affect WM and LTM?

The elaboration process can be distinguished from mere re-reading by the accompanying distributed patterns of fMRI activity in the brain. Whereas elaboration showed no benefit for WM, it did facilitate LTM performance for young but not old adults. Accordingly, in young but not old adults, the degree of neural separability of re-reading vs. elaboration was positively correlated with the elaboration benefit in LTM: Greater separation between the neural processes of reading and elaboration was associated with larger LTM benefits of elaboration across subjects (*Figure 7* C & D). The present results confirm prior studies showing evidence against a WM benefit of elaboration (Loaiza & Camos, 2016). However, our findings fail to provide experimental support for the conclusion from previous studies which found that higher WM performance on complex-span tasks was correlated with individuals' use of elaboration strategies such as imagery and sentence generation (Bailey, Dunlosky, & Hertzog, 2009; Bailey et al., 2008, 2011; Dunlosky & Kane, 2007). This discrepancy could be due to the present study using a simple-span paradigm and previous research relying on complex-span tasks. Alternatively, the correlation might not reflect a causal effect of elaboration on memory – rather, participants who have good memory have more information in memory to elaborate on.

8.5.3 Are refreshing and elaboration distinct processes?

If refreshing and elaboration are two labels for the same process, then the pattern of behavioral effects should be the same for WM and on LTM, and the patterns of brain activity supporting these processes should be indistinguishable. In the present study, in a combined mask of a priori brain regions from frontal, temporal, and parietal lobes, we found successful differentiation of brain activity associated with re-reading, refreshing and elaboration processes. This neural evidence supports the assumption that refreshing and elaboration are implemented with distinct neural processes.

Further, as discussed separately above, refreshing and elaboration resulted in distinct behavioral effects on tests of WM and LTM. Replicating our previous work (Bartsch et al., 2018), we found that elaboration, but not refreshing, improves episodic LTM, but not WM. Taken together, the neural and behavioral results replicate and extend previous findings by supporting a distinction of the refreshing and elaboration processes.

8.5.4 How do refreshing and elaboration contribute to age-related memory deficits?

In addition to the question of whether refreshing and elaboration are distinct processes that can facilitate memory, another goal of the present study was to investigate whether these processes and their impacts on memory are preserved in older adults. As in our young adult sample, the three processes of repeating, refreshing, and elaboration were neurally distinguishable in the predefined mask of frontal, parietal and temporal regions for a majority of the older adults (N = 22 of 27; Figure 8). The direct comparison of repeating vs. refreshing provided confirmatory evidence that, like young adults, the older adults engaged these processes differently, and that individuals with a larger degree of neural separability also had larger behavioral benefits of repeating over refreshing. In young adults, the *refreshing processing*

benefit also increased as the refreshing process became more similar to the repeating process in the brain. However, this relationship did not replicate for older adults. Instead, the *repeat processing benefit* was shown to decrease as these processes became more similar.

The interpretation we proposed for the young adults, of refreshing being most beneficial for memory when it appears in the brain to be like reading, does not appear to hold for older adults. Rather, our results indicate that refreshing is implemented differently, and less effectively, by older adults as compared to young adults. Future studies could account for these findings by, for example, instructing older adults to refresh an item specifically by "thinking of its visual representation". Such a manipulation would allow us to test whether implementing refreshing as a *simulated re-reading* of the memory items is most beneficial for WM, and it could also address whether a deficit in refreshing can be compensated for by teaching an effective strategy.

Similar to the young adults, refreshing had no benefit on LTM in older adults. This replicates the age-group specific findings of Johnson (2004), who also found no LTM benefit in old adults when comparing refreshing to re-reading. It was identified as an independent process, however, as refreshing was neurally separable from both re-reading and elaboration. As refreshing was not related to LTM performance, even in the young adults, we conclude that deficits in refreshing are not responsible for the LTM deficit in older adults either.

The results on elaboration show that the fMRI classifiers were able to differentiate mere re-reading from elaborating in the majority of the older adults (N = 20 out of 27; Figure 10). However, there was no LTM benefit of elaboration in older adults, whereas this effect was robust in the young group (*Figure 3*). We argue therefore that most of the older adults did perform some mental manipulation in the elaboration condition that was different from mere re-reading, but whatever it was did not affect their LTM performance. These results are in line with the

elaboration deficit hypothesis (Smith, 1980), showing that when having to generate their own elaborations (here mental images), older adults do not benefit in the same way as young adults do. Taken together, our results provide evidence that the LTM deficit of older adults might arise at least in part from a deficit in the process of elaboration. Future research might investigate whether age-related LTM deficits can be compensated by providing more precise elaboration instructions.

8.5.5 Conclusion

Our study revealed that the processes of repeated reading, refreshing, and elaboration are differentiable in brain activation patterns in both young and older adults. Refreshing was most effective when its neural pattern resembled that of repeated reading, which indicates the importance of reinstating the perceptual state of memoranda that are no longer physically present. Elaboration as a process can be neurally distinguished from mere reading. While it had no impact on WM, elaboration did improve episodic LTM for young adults and the size of the benefit was related to the neural separability of elaboration: The more differentiated elaboration was from re-reading, the more elaboration benefited LTM. In contrast to the young adults, older adults' episodic LTM did not benefit from elaboration, even though this process was neurally separable from reading. This suggests that older adults implemented a sub-optimal form of elaboration, and this may be a contributing factor to age-related.

8.6 Tables

Table 3 ROIs with their corresponding BAs and references of previous reporting in univariate analyses in the literature.

Label	sub region	BA	Labeled region reported in
frontal	inferior frontal	44,45,47	Johnson et al., 2005; Johnson, Mitchell, Raye, & Greene, 2004; Raye, Johnson, Mitchell, Greene, & Johnson, 2007; Raye, Mitchell, Reeder, Greene, & Johnson, 2008 Blumenfeld, 2006; Blumenfeld, Parks, & Yonelinas, 2010; Kim & Giovanello, 2011; Murray & Ranganath, 2007
	middle & superior frontal	4,6, 8,9,10,46	
parietal		3,7,40	Johnson et al., 2004; Kim & Giovanello, 2011; Murray & Ranganath, 2007; Raye, Johnson, Mitchell, Reeder, & Greene, 2002; Raye et al., 2007, 2008
fusiform		19, 37	Murray & Ranganath, 2007; Raye et al., 2008
parahippocampal		27, 28, 34, 35, 36	Kim & Giovanello, 2011 Murray & Ranganath, 2007
temporal	inferior	20	Blumenfeld et al., 2010;
	middle	21	Johnson et al., 2004;
	superior	22, 38	Kim & Giovanello, 2011; Murray & Ranganath, 2007

Table 4 The posterior effect estimates and their 95% HDRs of the generalized linear mixed model for binomial response variables for the immediate serial memory data.

	mean parameter on probit scale	**95% HDR**
(Intercept)	0.09	[-0.01, 0.19]
processing	**0.44**	**[0.38, 0.49]**
repeat/refresh	**-0.23**	**[-0.26, -0.19]**
elaboration	0.01	[-0.03, 0.05]
age	**-0.29**	**[-0.44, -0.15]**
processing * repeat/refresh	**-0.30**	**[-0.35, -0.26]**
processing * elaboration	-0.04	[-0.08, 0.01]
repeat/refresh * elaboration	-0.02	[-0.07, 0.02]
age * processing	0.01	[-0.07, 0.08]
age * repeat/refresh	**0.09**	**[0.04, 0.14]**
age * elaboration	0.05	[0.00, 0.10]
processing * repeat/refresh * elaboration	0.02	[-0.02, 0.08]
processing * repeat/refresh * age	-0.05	[-0.11, 0.02]
processing * elaboration * age	0.02	[-0.05, 0.08]
repeat/refresh * elaboration * age	-0.06	[-0.12, 0.00]
processing * repeat/refresh * elaboration * age	0.00	[-0.09, 0.08]

Note. Credible effects, defined as HDRs excluding zero, are printed in bold.

Table 5 The posterior effect estimates and their 95 % HDRs of the generalized linear mixed model for binomial response variables for the delayed memory data.

	mean parameter on probit scale	**95% HDR**
(Intercept)	**-0.29**	**[-0.38 -0.25]**
processing	**0.06**	**[0.01, 0.1]**
repeat/refresh	**-0.05**	**[-0.09, -0.01]**
elaboration	-0.04	[-0.09, 0]
age	**-0.06**	**[-0.11, -0.01]**
processing * repeat/refresh	-0.04	[-0.09, 0.02]
processing * elaboration	-0.01	[-0.07, 0.05]
repeat/refresh * elaboration	0.05	[-0.01, 0.11]
age * processing	0.03	[-0.03, 0.09]
age * repeat/refresh	-0.05	[-0.11, 0.01]
age * elaboration	**0.08**	**[0.01, 0.14]**
processing * repeat/refresh * elaboration	-0.03	[-0.12, 0.05]
processing * repeat/refresh * age	0.05	[-0.03, 0.14]
processing * elaboration * age	0.05	[-0.04, 0.13]
repeat/refresh * elaboration * age	-0.04	[-0.13, 0.05]
processing * repeat/refresh * elaboration * age	-0.04	[-0.15, 0.08]

Note. Credible effects, defined as HDRs excluding zero, are printed in bold.

8.7 Figures

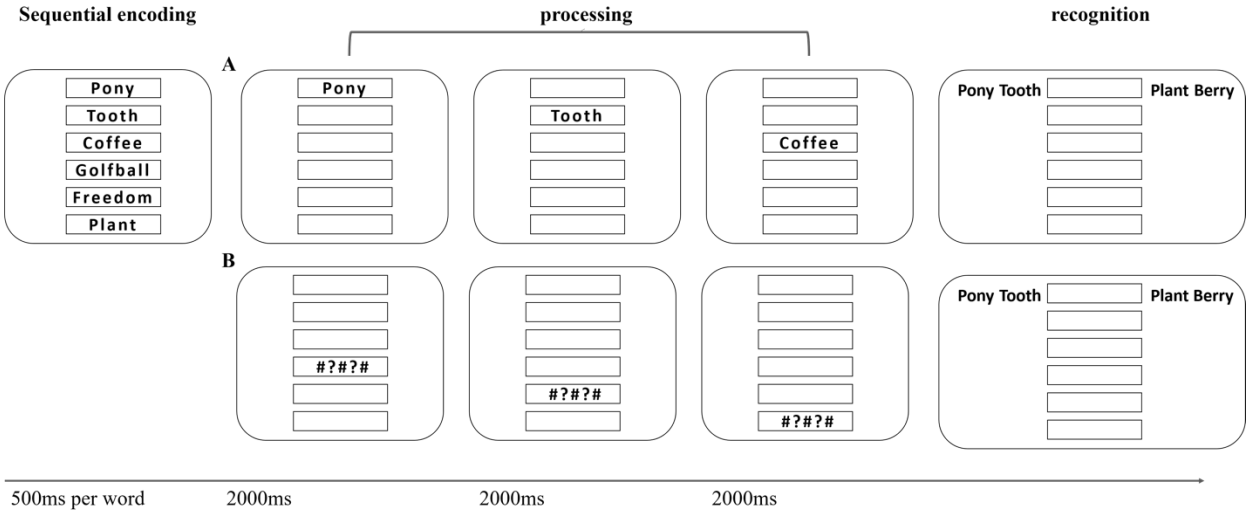


Figure 2. Illustration of the immediate memory paradigm. Subjects were shown a list of six words sequentially, followed by either the first (A) or second (B) triplet being processed according to the four experimental conditions. The trial ended with a recognition test in which each list item was tested in their order of presentation using a 4-alternative forced-choice procedure. (A) shows the repeat condition, which was preceded with the instruction to form a mental image in the elaboration condition. (B) shows the refresh condition, which was preceded with the instruction to form a mental image in the 4th condition (refreshing with elaboration).

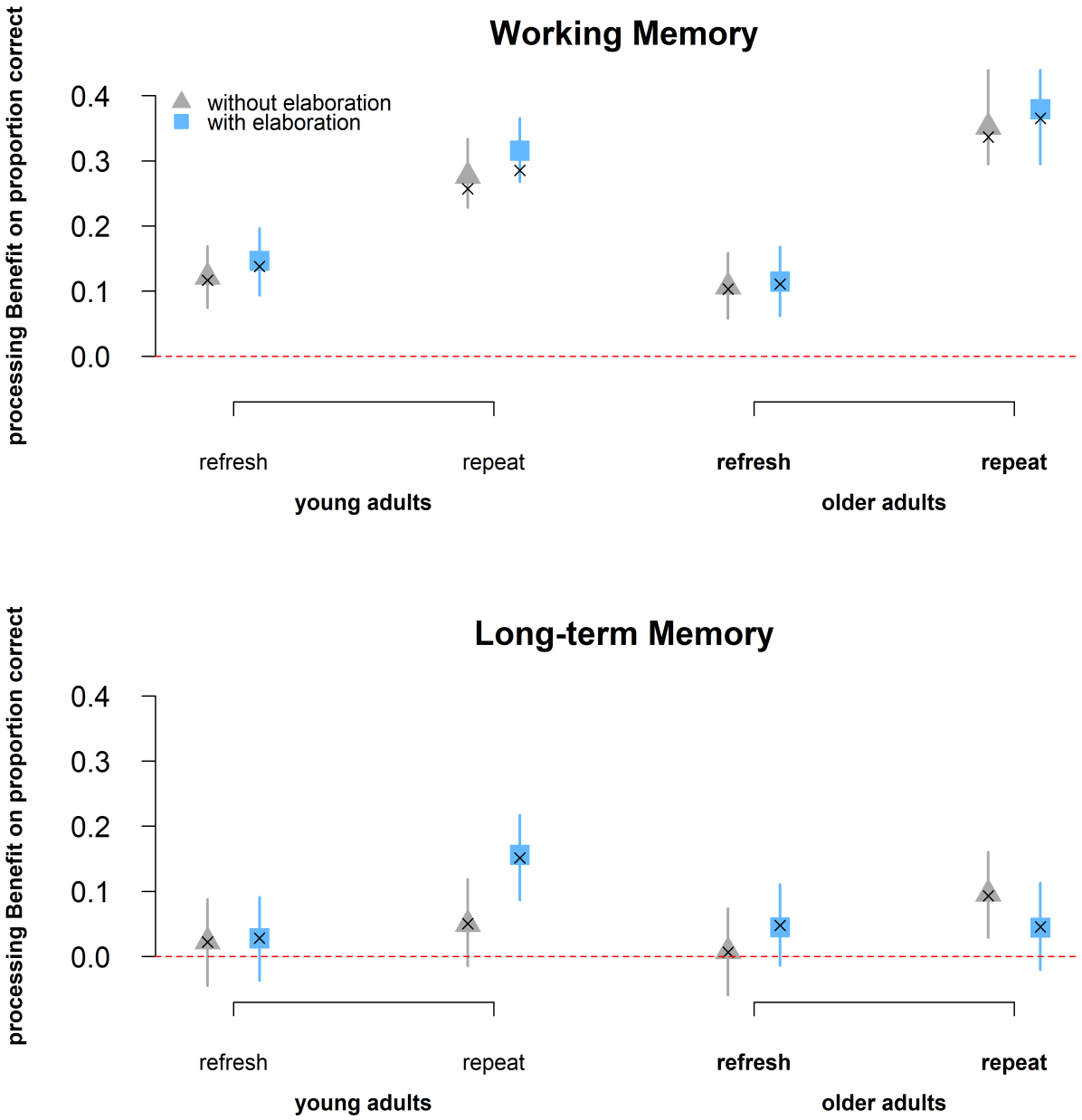


Figure 3 processing benefit in the WM (upper graph) and LTM (lower graph) task. The blue symbols and error bars represent estimated processing benefits and their 95% HDRs from the BGLMM for the conditions with elaboration, the grey symbols represent the same for the ones without elaboration. The crosses represent the observed data. Their overlap indicates that the model adequately describes the data. The red line represents the point of no difference in performance between the processed and the unprocessed triplet.

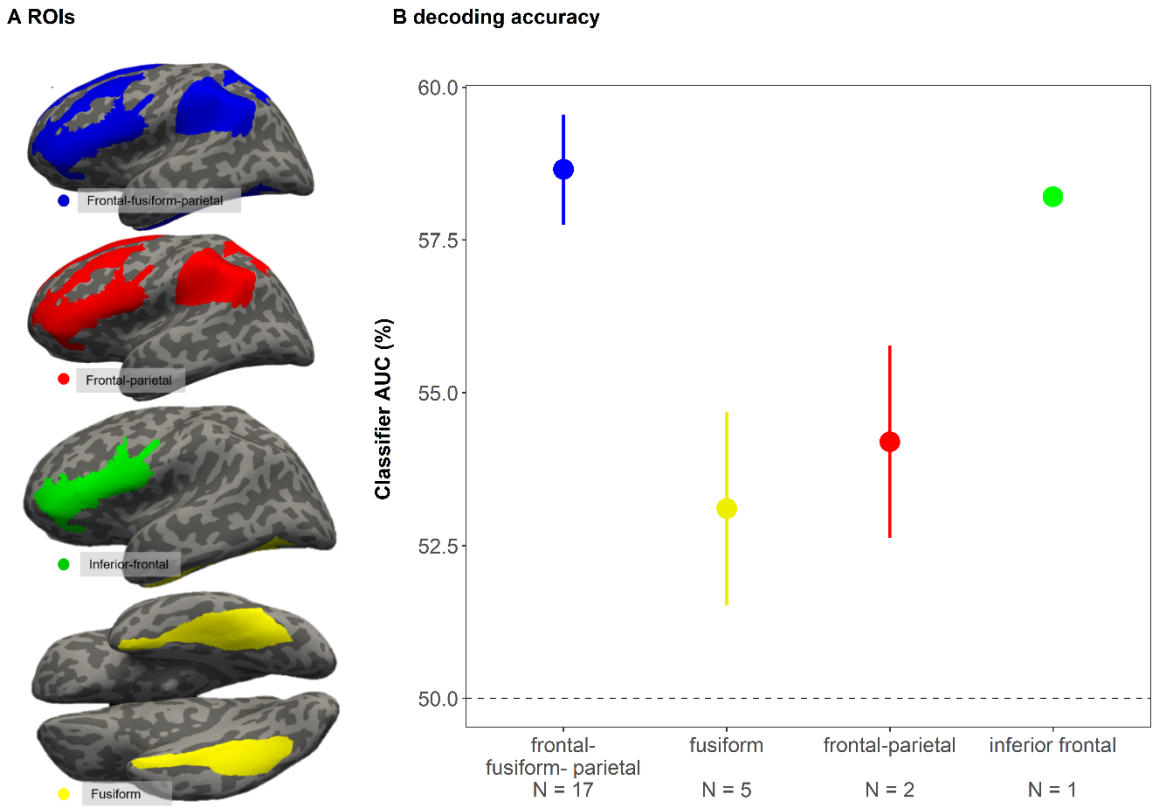


Figure 4 Classifier decodability across ROIs for Repeat vs. Refreshing vs. Elaboration in young adults. A: ROIs used for analysis. B: Decoding as indicated by the classifier AUC for the subjects in the frontal fusiform parietal mask or its subsets. The error bars indicate the standard error.

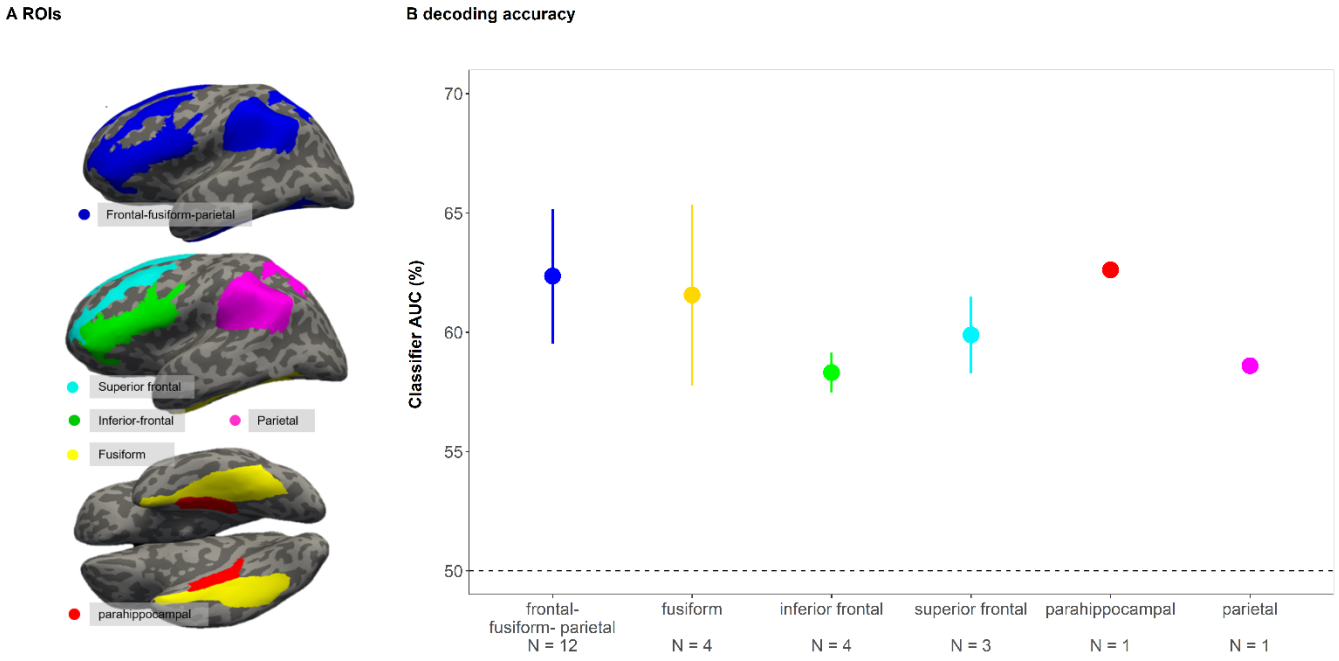


Figure 5 Classifier decodability across ROIs for Repeat vs. Refreshing for young adults. A: ROIs used for analysis. B: Decoding as indicated by the classifier AUC for each the subjects in the frontal fusiform parietal mask or its subsets. The error bars indicate the standard error.

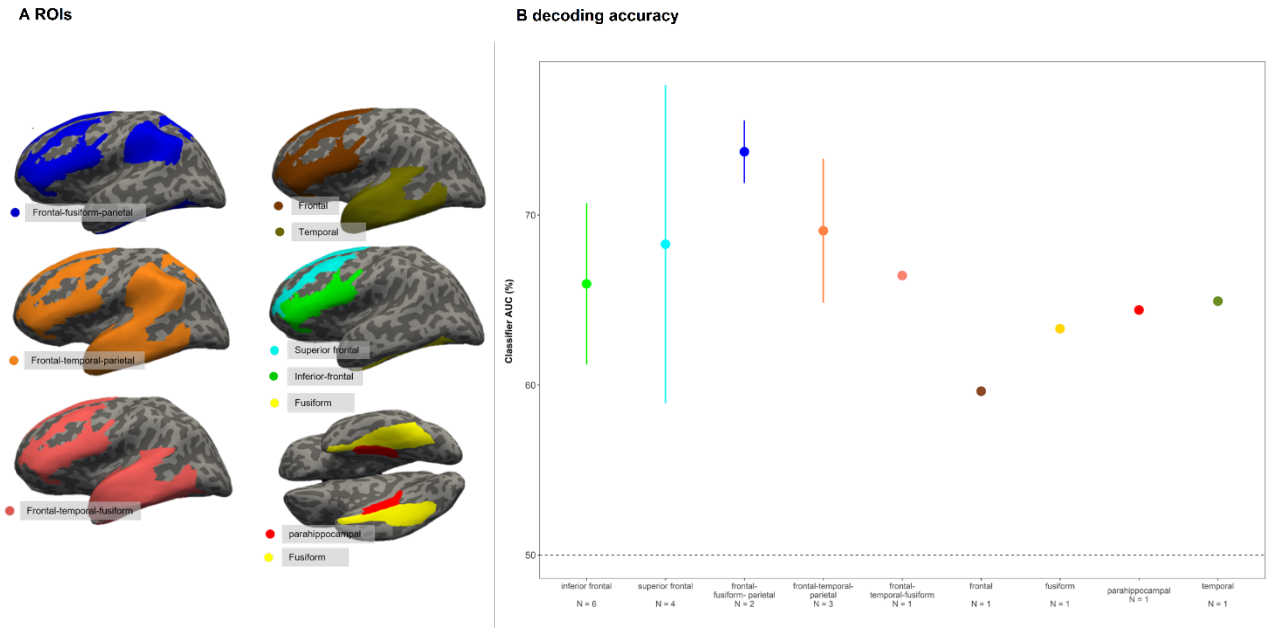


Figure 6 Classifier decodability across ROIs for Repeat vs. Elaboration for young adults. A: ROIs used for analysis. B: Decoding as indicated by the classifier AUC for each the subject in the frontal fusiform parietal mask or its subsets. The error bars indicate the standard error.

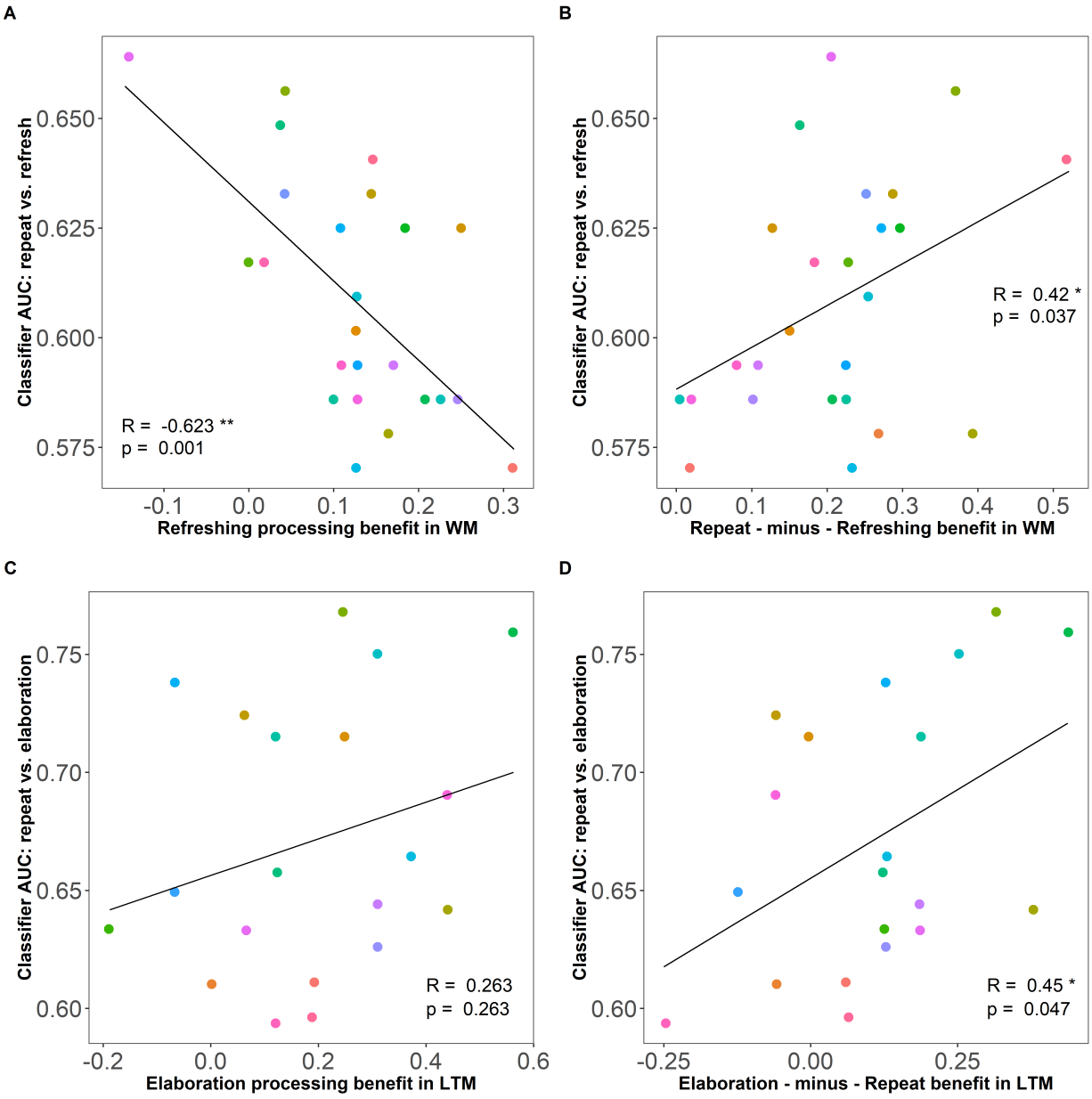


Figure 7 Correlation of MVPA classification (i.e. AUC) to memory performance. A: for the two-way problem of **Repeat vs. Refresh** in young adults to the subjects' behavioral refreshing processing benefit in working memory and to B: the subjects' behavioral repeat vs. refreshing benefit in working memory. Correlation of Classifier AUC for the two-way problem of **Repeat vs. Elaboration** C: to the subjects' behavioral elaboration vs. repeat benefit in long-term memory and D: to the subjects' behavioral elaboration vs. repeat benefit in long-term memory. Each point represents an individual subject.

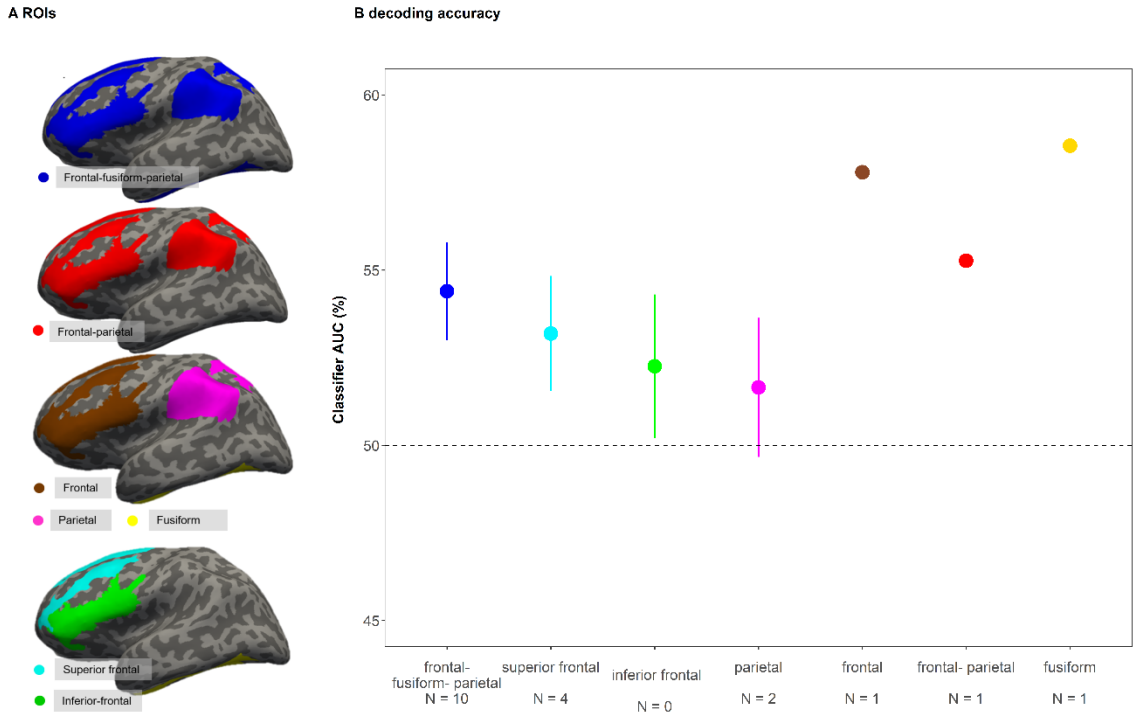


Figure 8 Classifier decodability across ROIs for the three-way problem Repeat vs. Refreshing vs. Elaboration in Older adults. A: ROIs used for analysis. B: Decoding as indicated by the classifier AUC for each the subjects in the frontal fusiform parietal mask or its subsets. The error bars indicate the standard error.

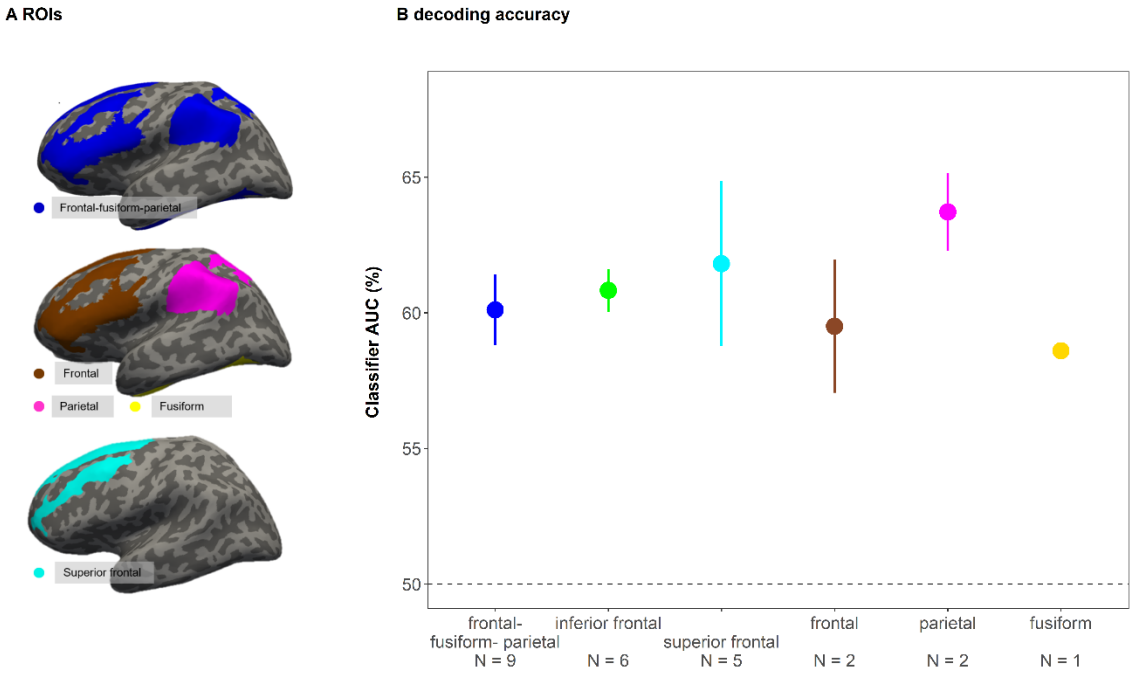


Figure 9 Classifier decodability across ROIs for Repeat vs. Refreshing in older adults. A: ROIs used for analysis. B: Decoding as indicated by the classifier AUC for each the subjects in the frontal fusiform parietal mask or its subsets. The error bars indicate the standard error.

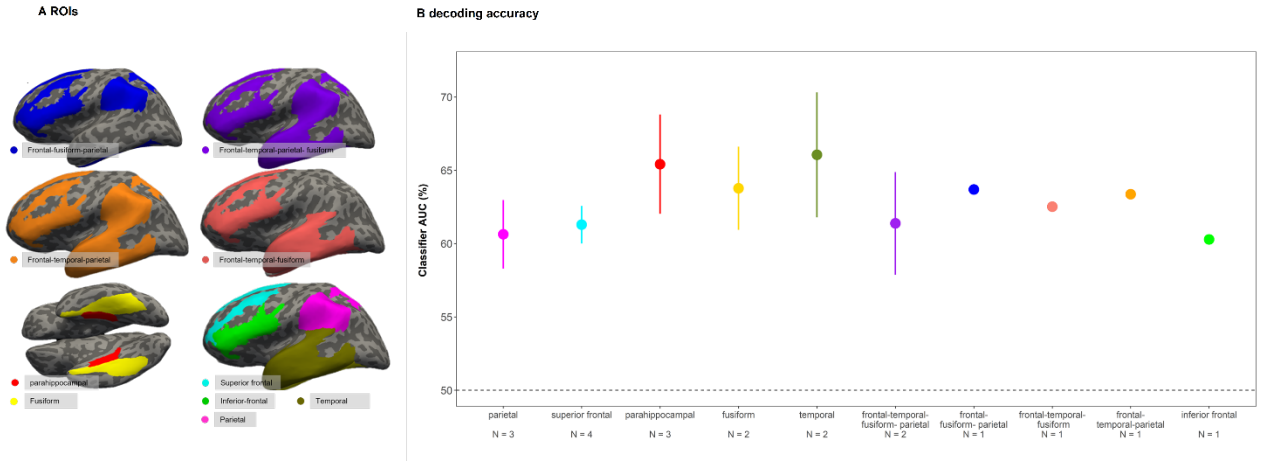


Figure 10 Classifier decodability across ROIs for Repeat vs. Elaboration in older adults. A: ROIs used for analysis. B: Decoding as indicated by the classifier AUC for each the subjects in the frontal fusiform parietal mask or its subsets. The error bars indicate the standard error.

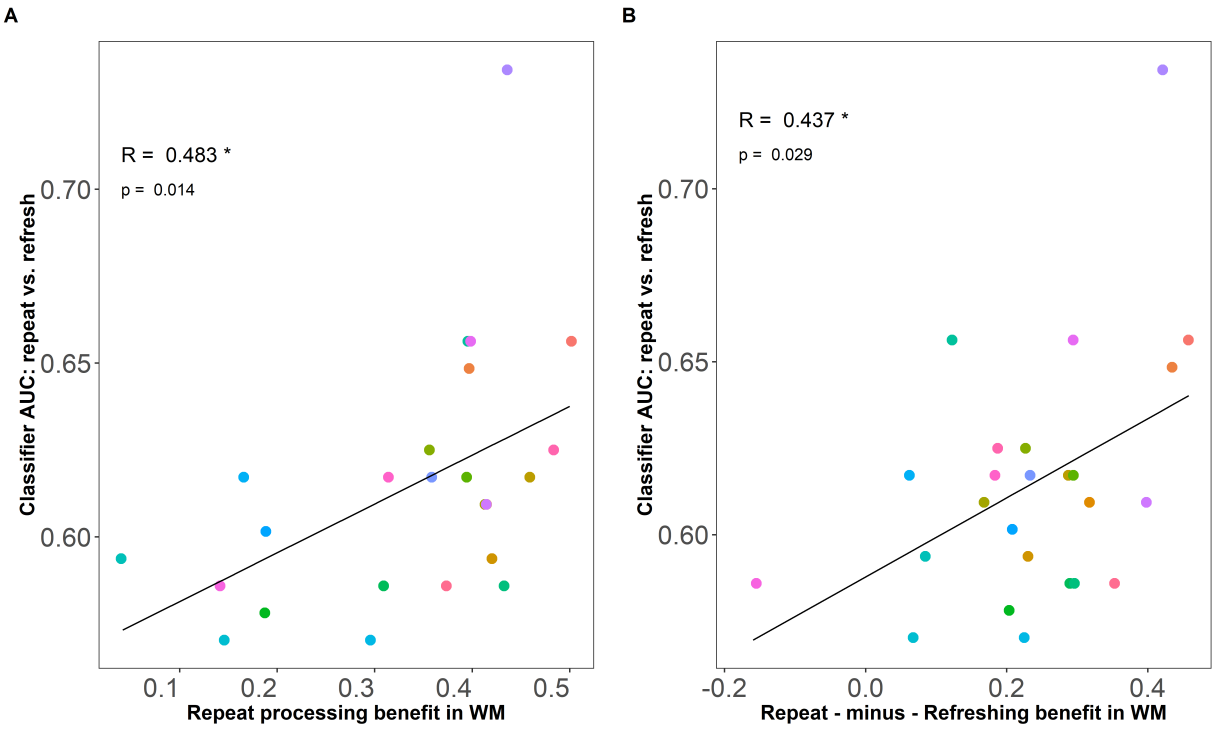


Figure 11 Correlation of MVPA classification (i.e. AUC) for the two-way problem of Repeat vs. Refresh in older adults to memory performance A: the subjects' behavioral repeat vs. refreshing benefit in working memory and B: the subjects' behavioral repeat processing benefit in working memory. Each point represents an individual.

9. Does limited working-memory capacity underlie age differences in associative long-term memory?

Lea M. Bartsch[1], Vanessa M. Loaiza[2] & Klaus Oberauer[1]

[1]University of Zurich and University Priority Program Dynamics of Healthy Aging

[2]University of Essex

Status: Resubmitted revisions to *Psychology and Aging* (August 28th, 2018)

Authors' contributions

LB reviewed the literature, designed the study, programmed the task, organized data collection, analyzed and interpreted the results and wrote the manuscript

VL designed the study, programmed the task and revised the manuscript

KO designed the research question, supervised and discussed LB's contributions, revised the manuscript

Author Note

We thank Jessica Grub, Gary Hoppeler and Danielle Pessach for helping with data collection. The data and the analysis scripts can be accessed on the Open Science Framework (https://osf.io/n8es5/). The project was funded by the University Priority Program Dynamics of Healthy Aging.

9.1 Abstract

Past research has consistently shown that episodic memory (EM) declines with adult age and, according to the associative-deficit hypothesis, the locus of this decline is binding difficulties. We investigated the importance of establishing and maintaining bindings in working memory (WM) for age differences in associative EM. In Experiment 1 we adapted the presentation rate of word pairs for each participant to achieve 67% correct responses during a WM test of bindings in young and older adults. EM for the pairs was tested thereafter in the same way as WM. Equating WM for bindings between young and older adults reduced, but did not fully eliminate, the associative EM deficit in the older adults. In Experiment 2 we varied the set size of word pairs in a WM test, retaining the mean presentation rates for each age group from Experiment 1. If a WM deficit at encoding causes the EM deficit in older adults, both WM and EM performance should decrease with increasing set size. Against this prediction, increasing set size did not affect EM. We conclude that reduced WM capacity does not cause the EM deficit of older adults. Rather, both WM and EM deficits are reflections of a common cause, which can be compensated for by longer encoding time.

9.2 Introduction

The long-term retention of episodes and events in episodic memory (EM) declines in older age (e.g., Hoyer & Verhaeghen, 2006; Naveh-Benjamin & Old, 2008; Zacks, Hasher, & Li, 2000). So far, the cause of this decline is still under debate. The proposed explanations include reduced speed of processing (Salthouse, 1996), reduced processing resources such us a limited working memory (WM) capacity (Craik & Byrd, 1982), and reduced inhibition (Zacks & Hasher, 1994). The age-related deficit in EM has been characterized as primarily a deficit of old adults in

building and retrieving relations (the associative deficit hypothesis, ADH, Naveh-Benjamin, 2000), whereas memory for individual components – referred to as item memory – is relatively intact in older age. The specific age-related decline in associative memory has been shown for various materials including word pairs, picture pairs, and face-name pairs (Bastin & Van Der Linden, 2005; Buchler, et al., 2011; Hara & Naveh-Benjamin, 2015). A meta-analysis evaluating 90 studies on the age-related associative deficit reports large effects sizes of age ($d_A \geq 0.80$) for verbal materials tested with a recognition test (Old & Naveh-Benjamin, 2008). Further, the meta-analysis showed larger age effects on associative than on item memory. Further the meta-analysis provided evidence that the size of the age-related associative deficit depends, among other variables, on the type of binding formed, with larger deficits for item-item compared to item-context bindings. Furthermore, the age-related associative deficit is larger for recall than recognition test formats (Old & Naveh-Benjamin, 2008). To have a clear characterization of EM decline in older age, it is a priority of cognitive aging researchers to isolate the causes for this disproportionate impairment in associative memory.

9.2.1 Age-related decline may be caused by a working memory deficit

Here we investigated whether WM plays a key role in causing associative deficits of EM in older adults (i.e., the *WM binding deficit hypothesis*). As an alternative, we consider the possibility that age-related deficits in WM and in EM are related through a common cause, such that inefficient encoding that impairs WM and EM alike (i.e., the *common cause hypothesis*). It should be noted that the terminology for relational information varies depending on the memory system. To clarify, for WM relational information is typically referred to as "bindings" whereas

for EM the term “associations” is used. We will refer to "bindings" as the general term, encompassing both bindings and associations.

The WM binding deficit hypothesis starts from the assumption that WM is needed to build and temporarily maintain new bindings, and that WM capacity is a limit on the maintenance of bindings (Oberauer, 2005). According to the WM binding deficit hypothesis, the capacity limit of WM could constrain the bindings formed in EM, and the age-related associative-memory deficit could be a consequence of older adults' reduced WM capacity (e.g., Chalfonte & Johnson, 1996; Hara & Naveh-Benjamin, 2015; Mitchell, Johnson, Raye, Mather, & Esposito, 2000; Park et al., 2002).

To justify the WM binding deficit hypothesis in the first instance, there should be evidence for an age-related deficit for maintaining bindings in WM that is similar to the one shown in EM. The evidence for this assumption is ambiguous. Some studies have provided evidence for an age-related binding deficit in WM (Borg, Leroy, Favre, Laurent, & Thomas-Antérion, 2011; Brown & Brockmole, 2010; Chalfonte & Johnson, 1996; Chen & Naveh-Benjamin, 2012; Cowan, Naveh-Benjamin, Kilb, & Saults, 2006; Fandakova, Shing, & Lindenberger, 2013; Mitchell et al., 2000; Oberauer, 2005) whereas others have found no evidence for it (Brockmole, Parra, Della Sala, & Logie, 2008; Brown, Niven, Logie, Rhodes, & Allen, 2017; Parra, Abrahams, Logie, & Della Sala, 2009; Peterson, Schmidt, & Naveh-Benjamin, 2017; Read, Rogers, & Wilson, 2016; Rhodes, Parra, & Logie, 2016).

Assuming that there is an age-related deficit in both maintaining WM bindings and remembering EM bindings, the question remains regarding how they are related: Does the WM binding deficit contribute causally to the associative EM deficit in old age? Support for this notion comes from work varying the study-test retention interval in a continuous recognition task,

revealing older adults' binding memory deficit over the short and long term (Chen & Naveh-Benjamin, 2012). However, this result is also consistent with the *common cause hypothesis* according to which binding deficits in WM and more long-term associative-memory deficits are both due to a more general inefficiency of encoding. Conclusive evidence for a causal role of impaired WM binding for the age-related associative deficits in EM has not yet been provided. The most convincing evidence for such a role would indicate that experimentally varying any potential age-related binding deficit in WM has strong consequences for the associative deficit in EM. More precisely, if the age-related binding deficit in WM were eliminated, then the corresponding associative deficit in EM should likewise disappear; similarly, if WM bindings were further impaired, then the associative deficit in EM should become larger.

Hara and Naveh-Benjamin (2015) indirectly tested this prediction by simulating older adults' EM associative deficit in young adults that had to perform a math task with varying difficulty while encoding name–face pairs. Their results showed that young adults' associative memory declined more than their item memory when engaging in a secondary task during encoding compared to full attention at encoding. This performance pattern mimics the older adults' associative memory deficit. The researchers concluded that the associative deficit in older age is caused by a reduction of their WM resources that may already be exhausted after the individual components are stored but before they are bound together.

One limitation of the study of Hara and Naveh-Benjamin (2015) is that their interpretation relies on an ordinal interaction: Binding memory of the young adults was already worse than their item memory at full attention; under divided attention, this difference increased. This interaction could simply arise because overall performance declines, thereby shifting performance into a more sensitive part of the measurement scale (i.e., further away from ceiling)

and consequently amplifying the contrast between item memory and binding memory. This ambiguity in ordinal interactions (Loftus, 1978) prevents any strong interpretation of the study of Hara and Naveh-Benjamin.

Here we take a different approach and test two predictions from the WM binding deficit hypothesis: First, if older adults' WM binding deficit is compensated by giving them more time for encoding the given bindings, this should also compensate the age difference in a subsequent test of EM for the same bindings. This prediction, however, also follows from the assumption that age-related encoding deficits are a common cause of older adults' impaired WM for bindings and their impaired EM for bindings. The second prediction can adjudicate between these two hypotheses: If increasing the number of items to remember (i.e., the memory set size) impairs the quality of bindings in WM, then increasing set size should likewise impair subsequent EM for the same information in both young and old adults alike. The two experiments of our study tested these two predictions.

9.2.2 Present Study

The goal of the present study was to investigate the importance of establishing and holding bindings in WM to age differences in retention of those bindings in EM. Accordingly, the two reported experiments investigated whether a WM deficit causes the age-related decline in EM. In Experiment 1 we investigated how equating memory for bindings in WM between young and older adults influences older adults' EM for the same bindings. We aimed to equate WM for bindings by adapting the presentation rate of the memoranda according to the subjects' ongoing performance on the WM task, particularly their retention of the bindings, as detailed below. If age-related differences in WM capacity cause the associative-memory deficit in EM,

then equating WM binding performance between young and older adults should eliminate the age-related deficit in EM binding. Experiment 2 aimed to test the same hypothesis through a second approach: If WM capacity limits the acquisition of bindings in EM, then increasing the load on WM (i.e., the memory set size) should impair binding memory in a WM test and also in a subsequent EM test for the same bindings.

9.2.3 Measuring Binding and Item Memory

In general, short-term relational recognition tasks require participants to retain bindings between each item (e.g., a word or an object) and another element, such as the item's context (e.g., locations on the screen in which they were presented) or another item (e.g., pairings of words with other words). During the test phase, participants are required to distinguish between the original pairings, recombined pairings, and pairs of new items. Older adults have exhibited more errors on such tasks requiring temporary bindings in WM compared to younger adults, and compared to tasks wherein only an item recognition decision, regardless of the item's bindings to other elements, is required (Fandakova, Shing, & Lindenberger, 2013; Oberauer, 2005). Researchers have subsequently varied these binding tasks to include more types of pairings and stimuli, and modified some details of the test format (De Simoni & von Bastian, 2018; Wilhelm, Hildebrandt, & Oberauer, 2013). For example, Wilhelm and colleagues (2013) presented pairings between two stimuli, such as words and digits, and randomly probed memory for the pairs, with one of the items of the pairing serving as a cue for either its correct match, a completely new item, or an intrusion of an item presented within the trial but not in its correct pairing. This work has provided evidence that a general binding factor represents a common

source of variance in typical WM tasks (e.g. complex span, Updating, Recall-1-back; Wilhelm et al., 2013).

Building on the WM binding task of Wilhelm et al. (2013), the WM task in the present study was an immediate memory test in which participants remembered arbitrary word pairs (e.g., *dog – tooth, tree – bottle*) and were tested with a three-alternative forced choice procedure. The test was designed to obtain separate measures of item and binding memory. Specifically, one item from each pair (e.g., *tooth*) was presented with options that included the original correct pairing (e.g., *dog*), a never-presented incorrect item (i.e., a new item; e.g., *book*), or an incorrect lure item that was presented in the trial but not in that pair (e.g., *tree*; see *Figure 12*). EM for the pairs was later tested with the same method used during the WM task. This paradigm allowed separate estimates of binding and item memory for both WM and EM in the same paradigm so that age differences could be investigated without confounding test differences with time of test.

In order to obtain estimates of binding and item memory from the responses in the above task, we applied multinomial process tree (MPT) models to the response frequencies (e.g., Buchner, Erdfelder, & Vaterrodt-Plünnecke, 1995). The structure of the MPT model is shown in *Figure 13*, and is equivalent to a measurement model reported in earlier work (i.e. the independence model, Jacoby, 1999; Jacoby, Debner, & Hay, 2001; see Cooper, Greve, & Henson, 2017 for recent evidence for this approach): The first branch represents whether or not the person correctly remembers the binding of the tested pair. If they remember the binding (with probability *Pb*), they can recollect the correct element previously paired with the cue. If they fail to remember the binding (with probability 1-*Pb*), they can still remember which items have been presented in the current trial (with probability *Pi*), In that case, they can guess between the two items that were in the current trial, leading to a correct response or to a lure response with equal

probability (g_b = 0.5). If they fail to recall the items in the trial (with probability 1-*Pi*), they guess among all three response options (correct, lure, new) with equal probability (g_i = 0.333).

The present implementation of the binding task, including the 3-alternative forced-choice set-up together with the MPT measurement model, allows purer estimates of binding and item memory compared to previous paradigms. More precisely, a pure measure of item memory is not achieved by instructing participants to only retain and report on single items – which were nevertheless presented in some context – because incidental encoding of bindings still affects performance (Jaswal & Logie, 2011; Prabhakaran, Narayanan, Zhao, & Gabriel, 2000; Postle, Awh, Serences, Sutterer, & D'Esposito, 2013; Reinitz & Hannigan, 2004; Treisman & Zhang, 2006). This evidence showing that binding memory contaminates many measures of item memory suggests that the aforementioned divergence regarding whether there is a specific age deficit for bindings may at least be partly due to an overestimation of the magnitude of an age difference in item memory. That is, if older adults have a true binding deficit and item memory may be affected by incidental encoding of bindings, then any age difference in item memory may be partly due to the binding deficit even though these are intended to be separate measures. Consequently, this could appear as a more symmetrical age difference between binding and item memory that does not accurately capture the true state of affairs. Thus, our relational recognition binding tasks, together with the MPT measurement model, may better identify the contributions of binding and item memory to performance, which is especially important given the mixed findings of the literature regarding an age-related binding deficit in WM.

In summary, the nature of the relational recognition task and the application of the MPT model allowed us to estimate relatively pure measures of binding and item memory for both WM and EM. Furthermore, the individual and ongoing adaptation of presentation rate of the pairs

based on a criterion of correct recollection of bindings in Experiment 1 allowed us to equate WM binding memory between age groups, and to use the resulting presentation rate for Experiment 2. These advantages of the study's design allowed us to distinguish whether equating binding memory in WM between age groups compensates for the age-related associative deficit in EM. Furthermore, if older adults' WM for bindings is impaired more than their WM for items, we expected that equating both age groups with regard to WM for bindings should lead to an age-related *benefit* for item memory (Old & Naveh-Benjamin, 2008).Finally, Experiment 2 utilized the presentation rates approximated in Experiment 1 to assess whether set size similarly impairs binding memory in WM and EM.

9.2.4 Analytic Approach

We used Bayesian statistical analyses, which have been recommended repeatedly for psychological research (e.g., Gallistel, 2009; Rouder, Speckman, Sun, Morey, & Iverson, 2009; Wagenmakers, 2007). Specifically, we implemented hierarchical Bayesian multinomial processing tree (MPT) models. MPT models estimate the probability of latent cognitive states on the basis of categorical data (Batchelder & Riefer, 1999). The hierarchical modeling framework accounts for participant heterogeneity by assuming that the individual parameters are drawn from a distribution describing the population. In this way, the model estimates parameters for each individual, as well as the mean and the dispersion of parameter values in the population. We fit hierarchical MPT models for each age group separately within the *TreeBUGS* Package (Heck, Arnold, & Arnold, 2018) for R (R Core Team, 2017), using the default uniform priors of the package, which are justified in the article by Heck and colleagues (2018). The MPT model of

Figure 13 was applied separately to the responses of each age group in the WM task, and to those of the EM task.

The TreeBUGS package yields Markov-Chain Monte-Carlo (MCMC) samples from the posterior probability distribution of the population mean of the MPT parameters (i.e., estimates of the mean *Pb* and *Pi* for each age group). By subtracting the posterior samples of the young from those of the old adults, we obtained a posterior distribution of the age difference in these parameters. We computed the means and the 95% credibility interval of these differences to assess the effect of age on the MPT parameters (Smith & Batchelder, 2010). The mean of the posterior provides a point-estimate of the effect size (i.e., the central tendency of the posterior difference). The 95% credibility interval gives the smallest range of parameter values over which 95% of the posterior probability is concentrated, and as such provides an assessment of the uncertainty of estimation (i.e., the dispersion of the posterior difference). For inference, we consult the proportion of the posterior probability density of a parameter difference that is larger or smaller than zero; this proportion provides an estimate of the posterior probability that the true effect is positive or negative, respectively.

9.3 Experiment 1

9.3.1 Method

9.3.1.1 Participants

We recruited 30 students (15 female) from the University of Zurich and 30 healthy older adults (15 female) from the Zurich community as participants[7]. They were compensated with either 15 Swiss Francs (about 15 USD) or partial course credit for the one-hour experiment. Cognitive functioning was screened with the MMSE (Mini-Mental Status Examination; Folstein, Folstein, & McHugh, 1975), indicating age-typical cognitive abilities in our sample of older adults (M = 28.92, SD = 1.07, range = 27 - 30). Table 6 shows the descriptive statistics and posterior distributions of the age effects of our sample. The evidence indicates fewer years of formal education in the older compared with the young adults. The older adults showed better performance than the young adults in a computerized vocabulary test (Mehrfachwahl-Wortschatz Test version B, Lehrl, 2005), consisting of 37 items in which participants are supposed to find an existing word between four similarly sounding non-words. The MWT-B is a marker test for crystallized intelligence. Hence, our sample of young and old adults show typical differences in education and measures of crystallized intelligence (Li et al., 2004).

9.3.1.2 Materials and Procedure

The stimuli were randomly drawn from a pool of 589 German concrete nouns for each participant. The nouns were between three and nine letters long and had a mean normalized

[7] We extended our initial sample of 20 young and 24 older adults during the revision of the manuscript. The use of Bayesian statistics allows for the continuation of sampling (Rouder, 2014; Schönbrodt, Wagenmakers, Zehetleitner, & Perugini, 2017).

lemma frequency of 24.76/million (drawn from the dlexdb.de lexical database). Pairs of nouns were created randomly for each participant to serve as the memoranda.

The sequence of an experimental trial is illustrated in *Figure 12*. After the sequential presentation of three word pairs, participants were tested on each pair in random order. Memory was tested with a three-alternative forced-choice task, also illustrated in *Figure 12*: For each probe word, participants selected the word paired with it from three response options: correct, new, and lure (word from another pair). The position of the options on the screen was random, and participants used the mouse to select among them at their own pace. In order to estimate the latent cognitive states of remembering the bindings and items we calculated the number of responses for each of the three response options (correct, lure, and new) for each individual.

The three to-be-remembered word pairs in each trial were sequentially presented from the top to the bottom of the screen. The presentation rate depended on the participants' current cumulative percent of correct binding responses: The adaptive algorithm was a variant of Kaernbach's (1991) weighted up-down algorithm that adjusted the presentation time per pair to achieve performance at 67% *correct* responses (i.e., choice of the correct pairing) for each participant. The algorithm continuously monitored average proportion correct over a moving window of the previous 10 trials. The presentation time for each pair to begin the experiment was set to 1000 ms and 1500 ms for the young and older adults, respectively. For older adults, the presentation rate increased by 200 ms if their moving average performance dropped below 67% correct, and decreased by 100 ms if performance exceeded the criterion. We initially tested 24 young adults with the exact same adaptation method, which unfortunately led to a mean accuracy level higher than we aimed for. We therefore decided to test a new group of young adults, reported here, with stricter adaptation parameters, which theoretically aimed for a 60%

criterion, but in practice reached the 67% criterion more closely. More specifically, the presentation rate increased by 180 ms if their moving average dropped below 60% *correct* responses, and decreased by 120 ms if performance exceeded the criterion. For both age groups, the maximum and minimum presentation rates were 5000 ms and 200 ms, respectively.

There were five trials of the WM task in each block. An unrelated distracter task followed each block, in which the participants had to indicate the correctness of presented math equations (e.g. 9 x 8 = 72) for 1 minute. After that followed a delayed cued recognition test in the same format as the immediate test, wherein the participants were probed again with one of the words from each of the 15 pairs from the previous block. The probes were presented in random order, and the probe word as well as the correct response option (i.e., the word actually paired with the probe) were the same as during the WM test. However, the new word among the response options was a completely new word to the experiment (i.e., not the same new word as in the WM test) and the lure word was a word from another pair (i.e., not the same lure word as in the WM test). As during the WM test, the position of the options on the screen was random, and participants used the mouse to select among them at their own pace. This method allowed us to measure binding memory in WM and EM in the exact same format. The experiment comprised five blocks in total.

9.3.2 Results

One older participant's presentation rate had reached the maximum (5000 ms) in the last block; this person was therefore excluded from further analysis, leaving data from 29 older and 30 young adults. For the analysis of the presentation rates as well as the performance during the

WM and EM tasks, we also excluded the first block, as we considered it as time for adaptation of the algorithm.

The Bayesian t-test to assess the evidence for the difference of the mean presentation rates per pair revealed decisive evidence for a difference between young (M = 657 ms, SD = 398) and older adults (M = 1724 ms, SD = 932), as the posterior density (PD) of the age effect lies entirely to the left of zero (PD: 0% < 0 < 100%).

The proportion of responses in each of the three response categories (correct, lure, and new items) can be found in Table 7. The critical analysis concerned whether adapting the presentation rate of the word pairs resulted in equated WM for bindings between young and older adults. For this analysis, we compared the age groups with respect to the item-memory and binding-memory parameters derived from the MPT model in *Figure 13*. **Fehler! Verweisquelle konnte nicht gefunden werden.** shows that the adaptation of presentation rates virtually eliminated the age difference in WM binding, as the posterior densities of the WM binding parameter of the young and older adults are overlapping. *Figure 15* depicts the posterior of the age-group difference, showing that the age effect in the mean WM binding parameter is concentrated around zero. We predicted greater item memory in the older adults than the young adults, as their item memory deficit is assumed to be less pronounced than their binding deficit (Mitchell et al., 2000). The difference in the mean WM item parameter supports this hypothesis, as the posterior density lies entirely to the left of zero.

After having ensured equated WM for bindings between age groups, we next examined whether the EM binding deficit was also eliminated. The lower panels of *Figure 14* and *Figure 15* show that the age difference in EM binding parameter still persisted (posterior mean of the age difference = 0.12, highest density interval (HDI) = [0.03, 0.21]). For item memory in EM,

the difference in the mean parameter reflects an approximately zero age difference. To ensure that this pattern of results cannot be explained by a mere testing effect (i.e., an advantage of retrieved over non-retrieved information; Rowling, 2014), we also conducted this analysis using EM performance conditionalized on correct WM binding memory. If the pattern is consistent between the former and the conditionalized analyses, then the age deficit in EM bindings is unlikely to be attributable to any differential retrieval practice that the pairs received in WM. The evidence for a remaining age difference in the EM binding parameter persisted in this analysis (see *Figure 16*).

9.3.3 Discussion

The goal of Experiment 1 was to investigate the importance of establishing and holding bindings in WM to age differences in EM bindings. We successfully equated WM for bindings between age groups by adapting the presentation rate of to-be-remembered word pairs. Our results show that this did not eliminate the EM deficit in old age. This implies that older adults' EM deficit is not entirely caused by a WM deficit at encoding. Nevertheless, by compensating the WM deficit, we substantially reduced the age-related EM deficit in comparison to previous studies (see the meta-analysis of age differences in EM with effect sizes of >.80 for verbal material and recognition tests, Old & Naveh-Benjamin, 2008). Therefore, the EM deficit could still be in part due to a binding deficit in WM, as the remaining binding deficit in EM was rather small. Alternatively, the results of Experiment 1 could be explained by the *common cause hypothesis:* Age-related deficits in WM and EM could be reflections of a common cause, which is partly compensated for by longer encoding time, leading to the reduction in EM binding deficits. For instance, both forms of memory might suffer from a similar age-related slowing of consolidation, the hypothetical process converting fragile, transient representations into more

stable memory representations (Chun & Potter, 1995; Jolicœur & Dell'Acqua, 1998; Wixted, 2004).

The goals of Experiment 2 were to test whether the EM associative deficit is in part caused by a binding deficit in WM, and to distinguish that hypothesis from the *common cause hypothesis*. According to the *WM deficit hypothesis*, interference between the word pairs in WM causes them to be represented less precisely or less robustly in older adults than in young adults, thereby leading to impaired EM representations. Accordingly, increasing the number of word pairs (i.e., set size) in WM should have a corresponding detrimental effect on EM. We would therefore predict that increasing memory set size leads to poorer performance not only in an immediate WM test but also in a subsequent test of EM. Moreover, because WM capacity declines in older age (Hale et al., 2011), we would predict an interaction between set size and age group on EM, such that older adults should show worse EM performance than young adults, especially as set size increases from a low value (at which both age groups' WM capacity is still sufficient to maintain all bindings well) to a value at which the WM of an average young person can still accommodate all bindings whereas the WM of older adults begins to struggle. The *common cause account* of the findings of Experiment 1, by contrast, predicts that the critical factor for EM performance is the time given to attend to and process the individual pairs, independent of how many other pairs are held in WM concurrently. In this case, increasing set size while keeping the presentation rate per pair constant should have no effect on EM binding. In order to adjudicate between these accounts, in Experiment 2 we varied set size of the tested pairs in WM while holding constant the presentation rate of the pairs at that for which WM for bindings of three pairs was equal between young and older adults in Experiment 1.

9.4 Experiment 2

9.4.1 Method

9.4.1.1 Participants

We recruited a new sample of 30 students (15 female) from the University of Zurich and 30 healthy older (15 female) adults from the community as participants. They were compensated with either 15 Swiss Francs (about 15 USD) or partial course credit for the one-hour experiment.

Cognitive functioning was screened with the MMSE (Folstein et al., 1975), indicating normal cognitive abilities in the sample of older adults (M = 28.82, SD = 1.47, range = 27 – 30). As evident by the posterior densities of the age differences in Table 1, the older adults had completed fewer years of formal education than the young adult and showed better performance in a computerized version of the MWT-B vocabulary test (Lehrl, 2005) than the young adults.

9.4.1.2 Materials and Procedure

The materials and procedure for Experiment 2 were very similar to Experiment 1. The principal differences were the following: Set size was varied across trials by presenting between 2 and 6 word pairs per trial. As in Experiment 1, the word pairs were sequentially presented from the top to the bottom of the screen. Furthermore, presentation rate was held constant within each age group at the mean presentation rates of the first 20 young and 24 older adults from Experiment 1, at which young and older adults showed equated WM for bindings for a set of three pairs (young = 710 ms and older adults = 1760 ms)[8]. As in Experiment 1, memory for each pair was probed immediately and after a delay. There was one trial of each set size per block, and seven blocks in the experiment.

[8] These presentation times were derived from the average presentation time per age group from the initial sample of 20 younger and 24 older adults. The mean presentation times reported for Experiment 1, include additional subjects and are therefore slightly different.

9.4.2 Results

We analyzed the number of correct, lure, and new item responses with the same hierarchical MPT model as in Experiment 1 using the *TreeBUGS* package in the R environment. We applied separate MPT models for WM and EM, as well as for each set size level and age group. The proportions of responses in the three response categories (correct, lure, and new items) are shown in *Figure 17*. The critical analysis concerned whether increasing interference in WM through increased set size decreases EM for bindings, and if so, whether that decrease was more pronounced in older than young adults even when using the presentation rates for which young and older adults showed equivalent WM binding at one of the lower set sizes (3 pairs) in Experiment 1. Figure 18 shows the posterior estimates for the main effect of set size, as well as the interaction effect of set size by age, for the parameters of the MPT models.

Unsurprisingly, increasing the number of to-be-remembered pairs in a trial reduced the WM binding performance for those pairs, represented by the change in parameter *Pb* (PD = 0% < 0 < 100%). Furthermore, as indicated by the interaction effect, the age-related difference varied with set size: young adults showed worse WM binding performance than older adults at set size 4, 5 and 6 (PD: 98.7% < 0 < 1.3%, PD = 89.6% < 0 < 10.4%, PD: 99% < 0 < 1%, respectively), but better WM binding performance than older adults at set size 2 (PD: 5.3% < 0 < 94.7%). For set size 3, the posterior for the age difference was centered on zero (PD: 41.3% < 0 < 58.7%), replicating the finding from Experiment 1 of approximately equivalent WM binding performance in both age groups at this set size with the given presentation rates.

The analysis of the parameters for item memory in WM revealed no effect of set size, neither for young (PD = 29.2% < 0 < 70.8%) nor for older adults (PD = 36.3% < 0 < 63.7%). For the main effect of age, 99.7% of the posterior density lay to the left side of zero, implying – as in

Experiment 1 – an age-related *benefit* for item memory in WM. This means that, after compensating for older adults' difficulty with maintaining bindings in WM, their item memory was better than that of young adults.

The analysis of the binding parameter for EM revealed evidence that the age difference was rather small, as the posterior density included considerable proportions on both sides of zero (PD: 16.5% < 0 < 83.5%). The critical analysis concerned whether set size affected EM in a similar way as it affected WM performance of young and old adults. There was, if anything, a very small main effect of set size, as the posterior density included considerable proportions on both sides of zero (PD: 86.2% < 0 < 13.8%). The difference in parameters between the age groups at each set size are shown in Figure 18C. There was evidence for an interaction between set size and age (PD: 2.1% < 0 < 97.9%), driven by better EM for bindings of young adults at set size 2 & 3 ($PD_{setsize2}$: 6.5% < 0 < 93.5%, $PD_{setsize3}$: 7.8% < 0 < 92.2%), whereas EM for bindings was equivalent between age groups at larger set sizes ($PD_{setsize4}$: 47.8% < 0 < 52.2%, $PD_{setsize5}$ = 36% < 0 < 64%, $PD_{setsize6}$: 64.8% < 0 < 35.2%).

For the parameter of item memory in EM, the analysis revealed no evidence for a main effect of set size (PD: 62.6% < 0 < 37.4%), nor a main effect of age (PD: 26.8% < 0 < 73.2%), and no evidence for an interaction between them (PD: 64% < 0 < 36%).

As in Experiment 1, we further ensured that the pattern of results of Experiment 2 was not attributable to a mere testing effect. To this end, we analyzed the EM performance also conditionalized on whether or not the pairs were correctly remembered during the WM test. The analysis confirmed the negligible age deficit in EM binding and item memory when performance was conditionalized on accurate WM binding (binding memory: PD: 62.6% < 0 < 37.4%, item memory: PD: 58.7% < 0 < 41.3%). Furthermore, the conditionalized analysis similarly showed

negligible set-size effects for binding memory and item memory (binding memory: PD: 2.1% < 0 < 97.9%, item memory: PD: 64% < 0 < 36%).

9.4.3 Discussion

To summarize, as a successful manipulation check, in Experiment 2 we replicated the equated binding performance in WM between young and older adults at set size three with the presentation rates from Experiment 1. Also, we replicated the substantive finding of Experiment 1, namely, that the retention of the bindings in EM was better in young than in older adults at set size three, despite the age-related compensation in WM. Furthermore, set size had the expected detrimental effect on WM bindings in young and older adults.

Despite its detrimental effect on both age groups' WM for bindings, increases in set size had no such effect on EM for either age group. Instead, the findings showed that the small age-related differences of EM bindings at set sizes 2 and 3 disappeared at the larger set sizes, as older adults' EM for bindings slightly increased at higher set sizes, leaving no evidence for an impact of age on bindings in EM.

From the *WM deficit hypothesis,* we predicted a set-size effect not only on WM but also on EM. In addition, we predicted an interaction between set size and age group in EM, such that older adults should show worse performance than young adults particularly at higher set sizes The above findings refute both predictions, decisively ruling out the *WM deficit hypothesis.* Our finding that old adults' WM deficit can be compensated for by giving them longer time for encoding and consolidating the memory pairs, and that this largely (Experiment 1) or entirely (Experiment 2) compensated for their EM deficit, is better explained by the *common cause hypothesis.* Specifically, older adults might be slower in consolidating information in both WM

and EM, and this slowing is partially responsible for their reduced binding ability in tests of WM as well as EM.

Experiment 2 yielded one unexpected effect: The set-size effect on WM bindings was larger for young adults, resulting in worse WM performance relative to older adults at larger set sizes, and somewhat better performance at the smallest set size. This result is surprising in light of a recent study by Read (2016) showing that increases in set size similarly impaired feature-location and feature-feature bindings in younger and older adults. That said, our results are in line with findings from Boujut & Clarys (2016). We can only offer a speculative post-hoc explanation for this interaction of set size with age: We compensated the age-related WM binding deficit by giving older adults substantially more time for encoding each pair. We tentatively concluded that this time is used to consolidate bindings better in both WM and EM. Perhaps the longer presentation time is used primarily for establishing better memory representations in EM, with relatively little effect on WM. In addition to improving delayed memory, better EM representations could also assist performance in the immediate test (intended to measure WM). As a consequence, older adults' performance on the WM test would rely more strongly on EM than that of younger adults. As EM is not affected by set size, this would result in a flatter set-size effect in the WM-test performance of older compared to younger adults.

9.5 General Discussion

The goal of the present study was to investigate the importance of maintaining bindings in WM for age-related EM deficits, especially the disproportionate associative deficit in older age. Using a novel paradigm that adapted the presentation rate of word pairs for young and older adults, we equated WM for bindings and subsequently observed a small but persistent EM binding deficit in older adults in Experiment 1. Further, the results of Experiment 2 were

incompatible with the *WM deficit hypothesis*, suggesting instead that inefficiency at encoding or during consolidation of memory traces may cause a more general age-related deficit in retaining bindings in WM and EM alike, in line with the *common cause hypothesis*.

The finding that memory set size, although strongly affecting WM performance, had no effect on subsequent EM for the same information for either age group contradicts the *WM deficit hypothesis* that has been advanced in previous work. For example, Hara & Naveh-Benjamin (2015) simulated the age-related associative deficit in EM by having young adults encode materials under divided attention, and they interpreted their result as consistent with the *WM deficit hypothesis*, such that an associative deficit in WM causes EM associative deficits in older adults. The present investigation questions this claim and suggests an alternative explanation: Instead of simulating a WM deficit, the divided attention manipulation reduced the time available for encoding, and therefore impaired young adults' EM to a similar extent as a naturally occurring encoding deficit of old adults.

Although the current results rule out the hypothesis that WM binding deficits cause EM deficits in older adults, they leave us with a new question: Which process did older adults in our experiments engage during their longer encoding time to reduce their deficit in both the WM and the EM tests? One possibility could be that older adults invested the increased encoding time to use (more) normatively effective strategies, such as elaboration. Findings from Bailey, Dunlosky, and Hertzog (2009) speak against this notion: These researchers showed that young and older adults report a similar prevalence of normatively effective strategies during WM tasks. In contrast, measures of processing speed accounted for a substantial proportion of the age-related variance in WM performance.

Our findings are consistent with the general slowing hypothesis (Salthouse, 1996; for an overview see Hartley, 2006), which proposes reduced processing speed to account for age-related differences in cognitive functions. The general slowing hypothesis emerged from consistent observations that older participants show longer reaction times to respond to stimuli, which supposedly represents slowing of perceptual, motor, and cognitive processes. Although slowing as common cause for age-related deficits in many tasks is attractive for its parsimony, it has long been debated what actually causes the phenomenon. For example, one could interpret the present results as consistent with the slowing of consolidation, such that older adults differentially struggle to create stable memory representations and require more time to do so compared to younger adults, thereby causing binding deficits in WM and EM. However, the current study cannot dissociate whether general cognitive slowing or a more specific slowing of processes such as consolidation cause the age-related binding deficit.

Furthermore, although our results are in line with the general slowing hypothesis, they do not strongly support this interpretation because other interpretations are just as plausible. For instance, it has been shown that older adults suffer from more neural noise, and therefore create less distinctive representations between successively presented pairs (Noack, Lövdén, & Lindenberger, 2012). As a result, the pairs are encoded with more overlapping representations, and at retrieval, the probe cues other words in the trial in addition to the actual target. This would cause binding memory impairments while leaving intact memory for items. For the current study, longer processing time could have led to more distinct, less noisy representations given the greater temporal separation of the pairs. This would result in better distinctiveness of the material, thereby reducing the binding deficit of the older adults. At larger set sizes the additional time may be particularly useful to engage in differential encoding (i.e., forming representations of the

word pairs in the memory set that emphasize the differences between them). Accordingly, what may appear at first glance as a general slowing deficit could instead reflect more time to engage in specific processes that may be deficient in older age, such as greater use of normatively effective strategies, consolidation of traces into stable representations, and reduction of representational overlap.

The current research is also relevant to previous work that has considered variation of presentation time to examine age deficits in WM. For example, Oberauer and Kliegl (2001) applied an adaptive algorithm (Kaernbach, 1991) similar to our Experiment 1 to vary presentation rates for young and older subjects in a WM updating task. They showed that WM capacity limits of the old adults could not be fully compensated by increasing encoding and updating times, as the young adults benefited from longer times too, and reached a higher asymptotic performance level with increasing time for each updating step. These findings indicate that age deficits in binding cannot be solely attributed to slower encoding or consolidation. Other work has similarly tried to compensate for the age-related memory deficit by increasing encoding time but has not managed to fully do so. A study by Sander and colleagues (Sander, Werkle-Bergner, & Lindenberger, 2011) showed WM performance of older adults increased with longer presentation rates; nevertheless, the older adults did not reach the level of the young adults' performance. The failure to fully compensate the age-related WM deficit could have occurred because they chose fixed longer presentation rates for older compared to younger adults, rather than adaptively varying presentation rates as in the current study. The choice of the presentation rate for older adults might just not have been slow enough to fully compensate their WM deficit. Two further recent studies investigated the effects of encoding time on WM bindings in young and old adults: Rhodes et al. (2016) found no

differential effect of longer (2500ms) compared to shorter (900ms) encoding time on WM binding performance of older adults. Similarly, Brown et al. (2017) found that although older adults profited more than young adults from a longer encoding time overall, age-related visual binding deficits in WM persisted at both short (900 ms) and longer (1500 ms) encoding durations. The findings of these studies do not conflict with our results as these studies' procedures allowed both age groups a fixed amount of additional time at encoding, which also permits the young adults to improve their bindings in WM. This is different to our approach here, where we calibrated the encoding time to *compensate* the older adults' lower performance. Another difference between the present study and that of Brown et al. (2017) could be the use of different stimuli (i.e., binding of shape and color vs. pairs of unrelated words): meta-analytic evidence suggests the age-related associative deficit is smaller for verbal compared with visuospatial materials (Old & Naveh-Benjamin, 2008). Accordingly, visuospatial memoranda may require an even larger adjustment of encoding time for older adults to compensate for their relatively larger binding deficit.

In summary, the present study tested the causal role of WM for the age-related binding deficit in EM. The results ruled out the *WM deficit hypothesis* that asserts that the binding deficit is due to a deficit to establish and maintain bindings in WM. Instead, the evidence was congruent with a common cause of both deficits. One plausible candidate for this common cause lies in less efficient encoding and consolidation processes in older age.

Table 6 Sample Description (means (and standard deviations)) of Experiment 1 and 2

Experiment	Age Group	Age	years of education	vocabulary
1	Younger	24.06 (3.77)	14.70 (3.07)	77.07 (13.22)
	Older	71.26 (3.98)	13.81 (3.40)	86.29 (3.43)
	$PD_{age\text{-}effect}$	-	16.9% < 0 < 83.1%	99.9% < 0 < 0.1%
2	Younger	24.86 (2.88)	15.84 (3.08)	77.90 (14.26)
	Older	71.24 (3.71)	13.66 (4.54)	85.27 (6.95)
	$PD_{age\text{-}effect}$	-	3.1% < 0 < 96.9%	98.7 % < 0 < 1.3%

Note. The posterior density (PD) of the age effects. Zero represents the point of no age differences, and the percentages indicate how much of the estimated effect's posterior distribution lies below and above 0. Values below 0 reflect an advantage of older adults whereas positive values indicate a younger adults advantage.

Table 7 Percent of responses per category in Experiment 1.

Memory test	Age Group	Correct	Lure	New
WM	Younger	66.47 (4.36)	20.97 (4.11)	12.56 (2.88)
	Older	69.14 (7.04)	23.71 (7.4)	7.14 (4.41)
EM	Younger	61.81 (9.7)	25.61 (7.18)	12.58 (4.33)
	Older	55.43 (8.94)	31.71 (8.57)	12.86 (7.13)

Note. The standard deviation is marked in parentheses.

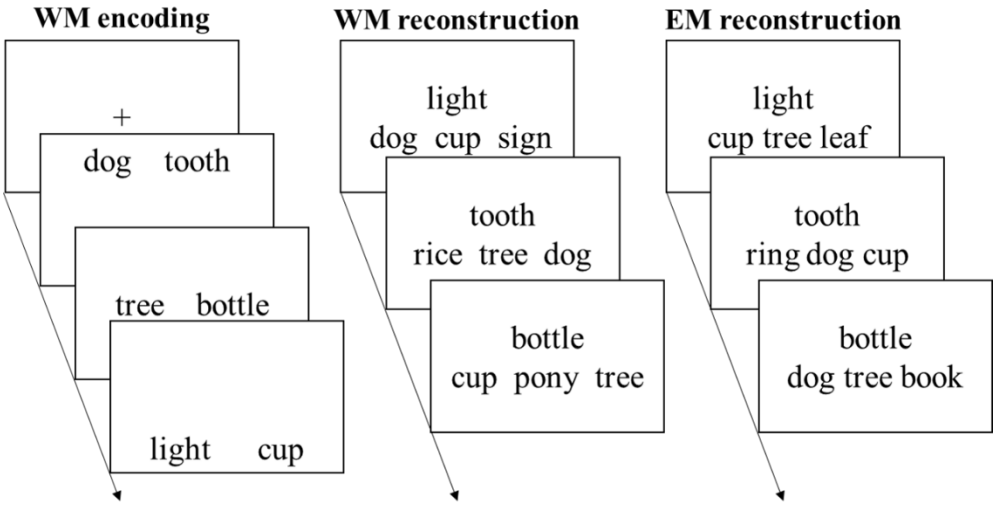


Figure 12 A representation of the paradigm used in Experiments 1 and 2.

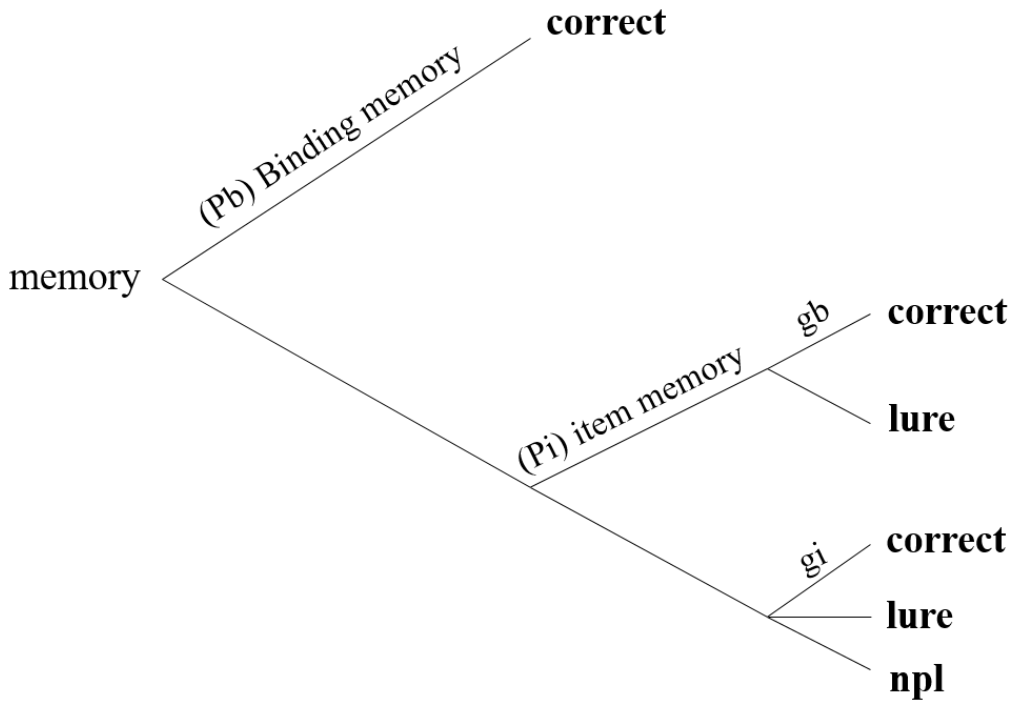


Figure 13 Multinomial-process tree (MPT) model for memory of bindings in Experiments 1 and 2. See section *Measuring Binding and Item Memory* for details.

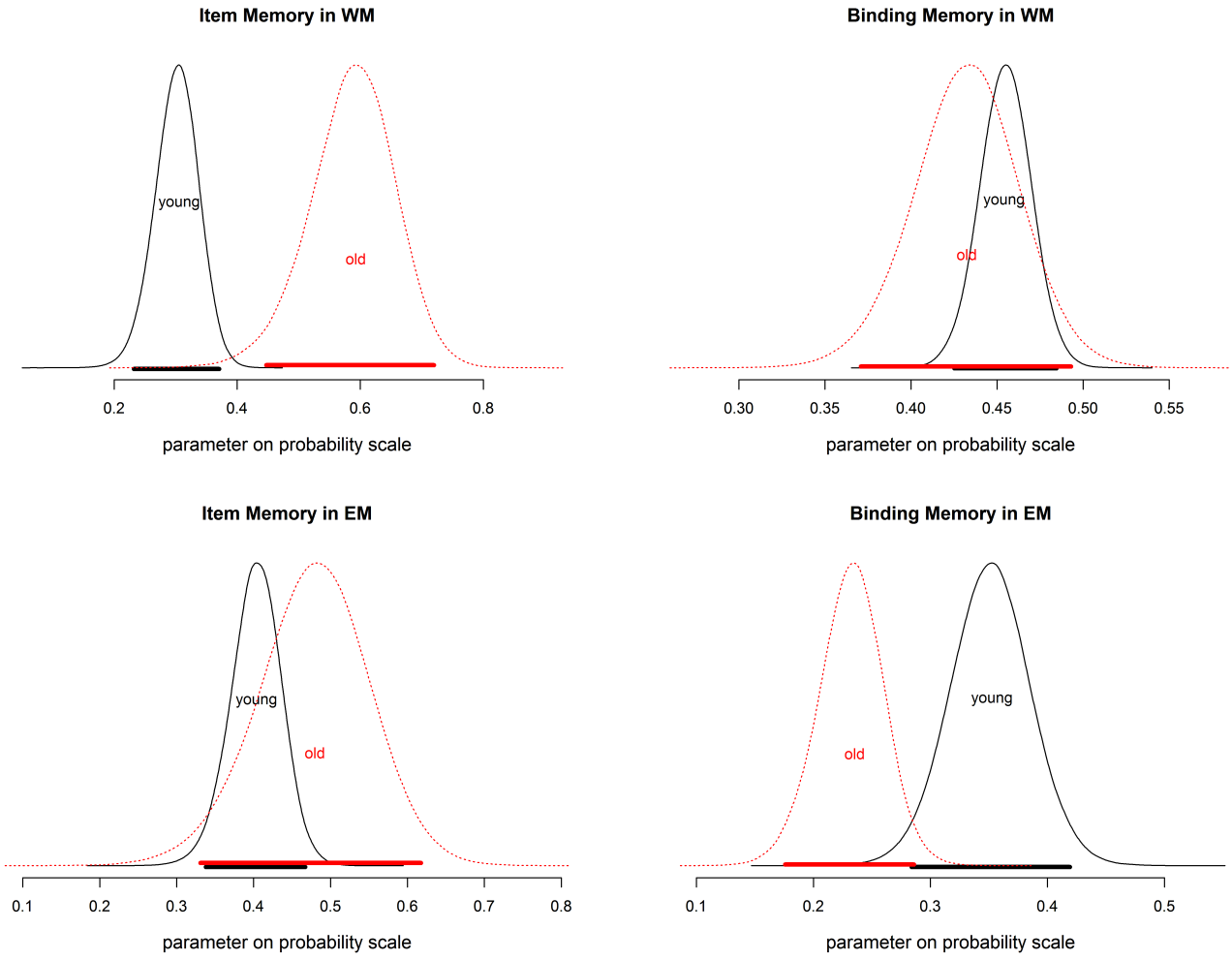


Figure 14 Posterior distributions of the parameters of the MPTs for young and older adults of Experiment 1. The horizontal lines represent the 95% highest density intervals.

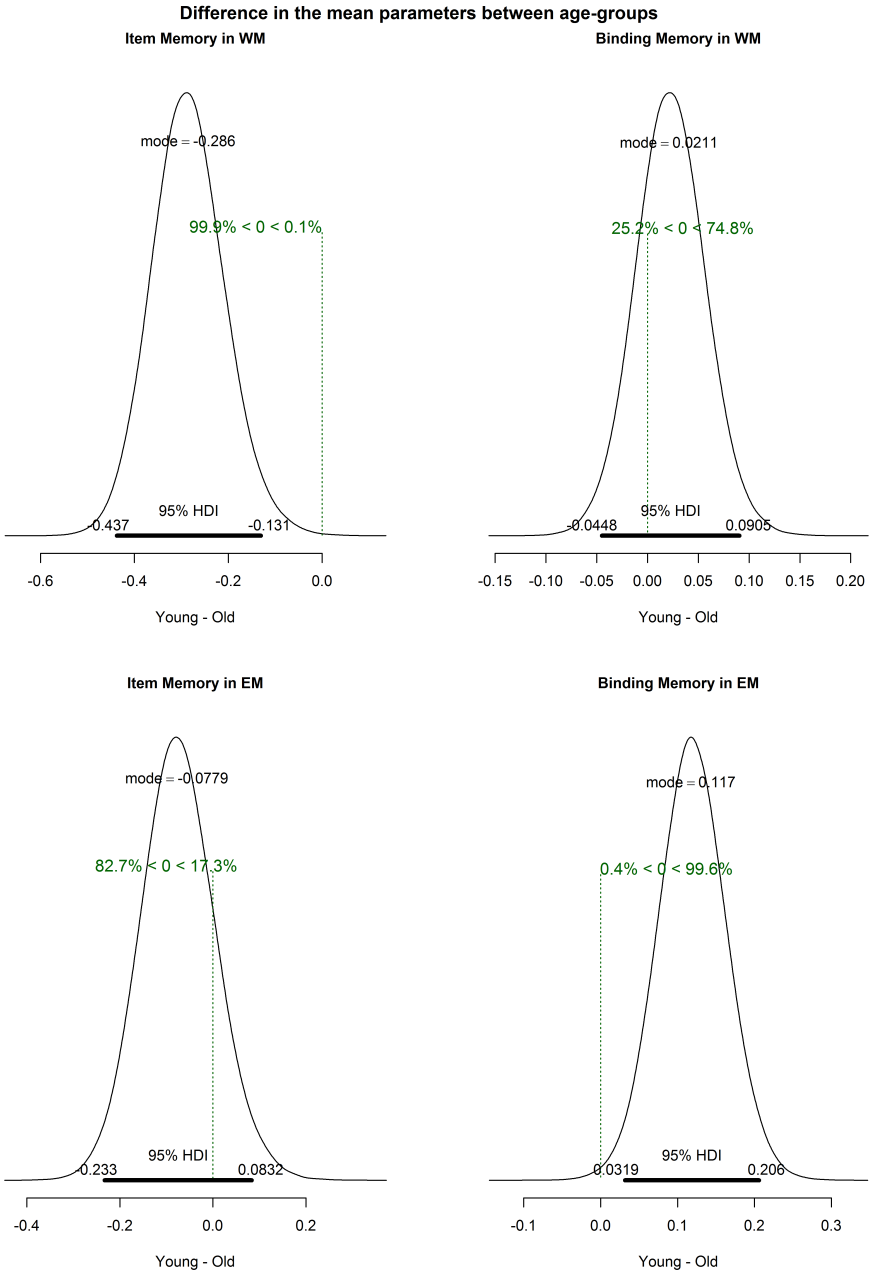


Figure 15 Posterior distributions of differences of mean the parameters between the age groups of Experiment 1. The mode with its respective highest density intervals reflect the effect size of any age difference. The dotted line indicates the point of no age differences, and the percentages indicates the credibility interval of the difference. Values below 0 reflect an advantage for the older adults.

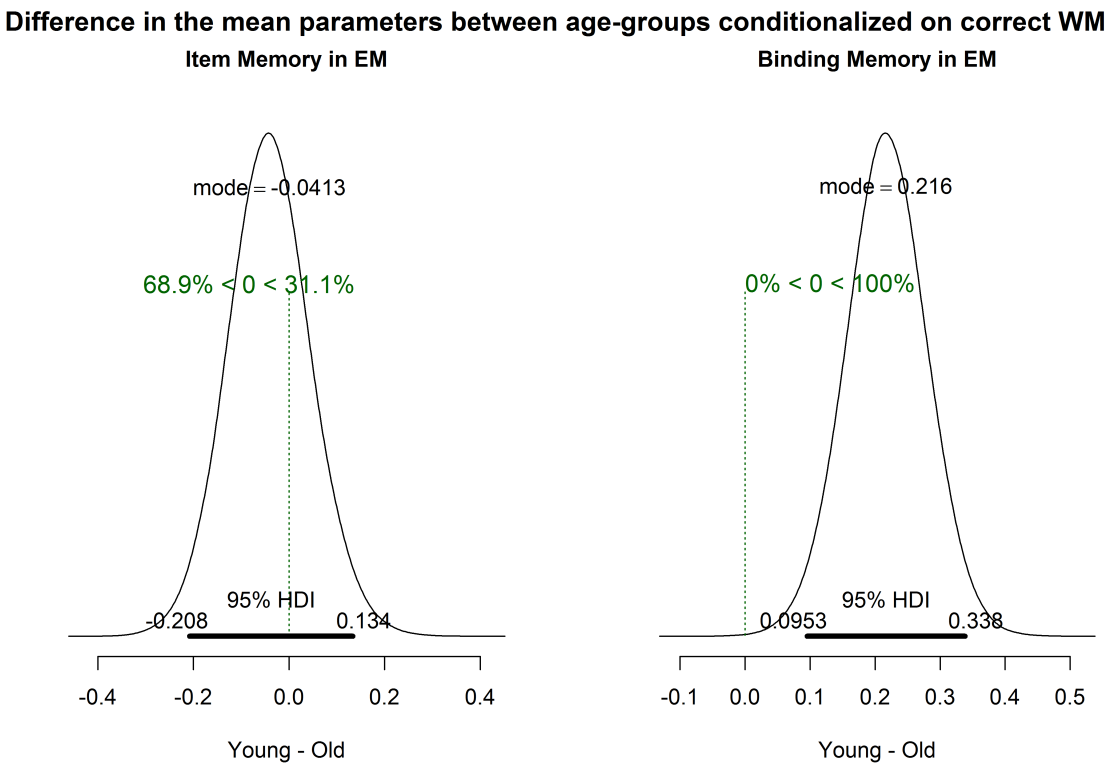


Figure 16 Posterior distributions of differences of mean the parameters between the age groups conditionalized on correct binding memory in WM of Experiment 1. The dotted line indicates the point of no age differences, and the percentages indicates the credibility interval of the difference. Values below 0 reflect an advantage for the older adults.

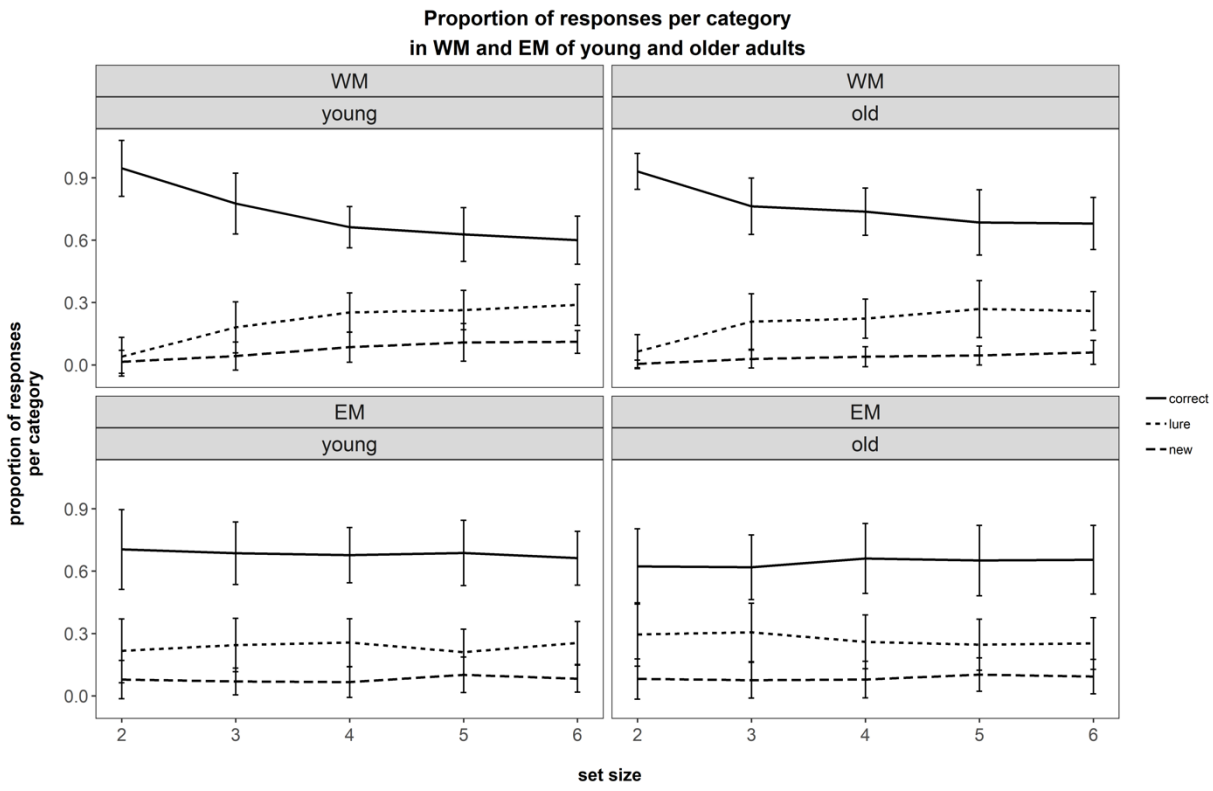


Figure 17 Proportion of responses per category in WM and EM of young and older adults in Experiment 2. The error bars represent the standard deviation.

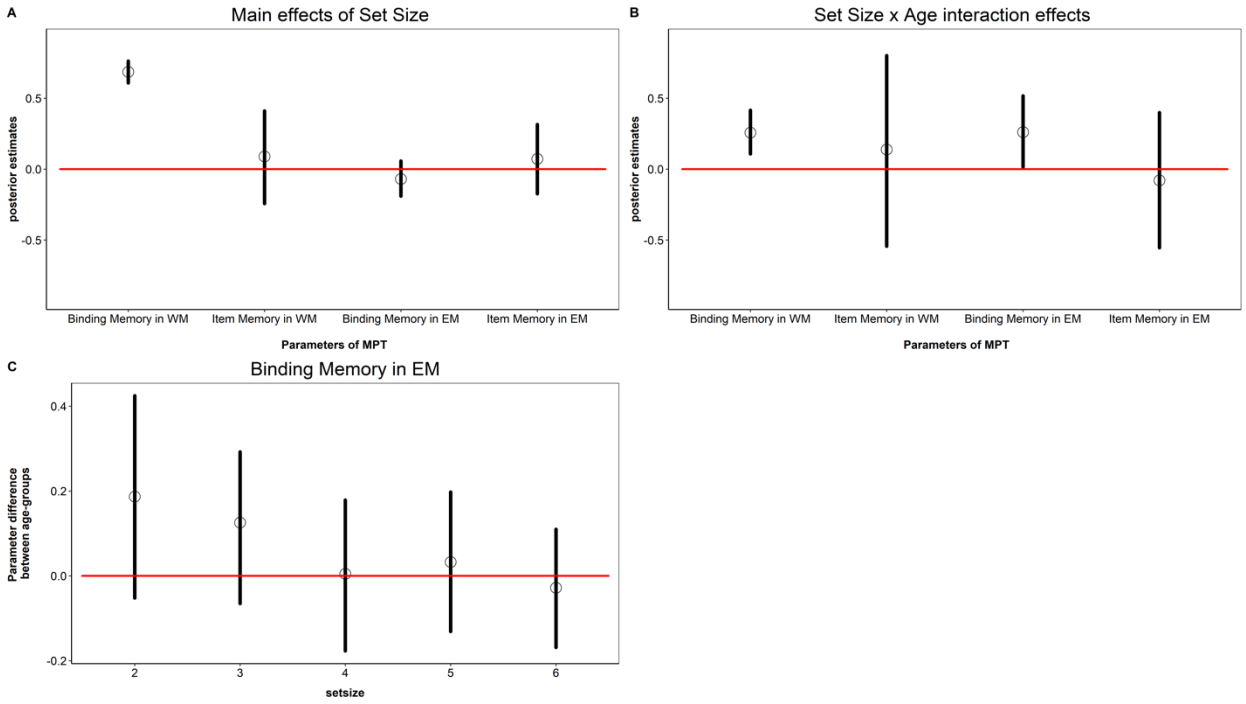


Figure 18 (A) Posterior estimates of the main effects of Set Size and (B) Posterior estimates of the interaction effects of set size with age-group of Experiment 2. Values above zero represent a stronger effect for young than for old adults. (C) Difference in posterior estimates of the parameter for binding memory in EM between the age groups per set size. Values above zero reflect an advantage for young adults. The red line characterizes the point of no evidence for an effect. The error bars represent the highest density regions.

10. Acknowledgements

I would first like to express the deepest gratitude to my thesis advisor Prof. Klaus Oberauer, for his constant advice, inspiration and guidance, for his open door at any time and for giving me the opportunity to do the work I appreciate so much.

I would also like to thank Prof. Yee Lee Shing who as a member of the examining committee reviewed this work. Thank you very much for the inspiring discussions that contributed to this work.

Furthermore, I had the privilege to spend three months at the University of Texas at Austin in the awesome *LewPeaLab*, were I was supported exceptionally by Prof. Jarrod Lewis-Peacock in conducting the analyses for my second paper. Thank you Ethan, Gus, Hyojeong, Jarrod, Remy, Seth, Steph, and Tracy for all your help and awesome Texas-experiences!

I would like to express my gratitude to the University Research Priority Program *Dynamics of Healthy Aging* and the Swiss Academy of Humanities and Social Sciences for their financial support of my co-authors and me.

I would like to show my greatest appreciation to Carla, Julia, Michel, Marcel, Mirko, and Tuule for sharing an office, for their continuous support. Special thanks also to all present and former members of the Cognitive Psychology Unit – again in alphabetical order: Alessandra, Alodie, Andra, Anke, Clara, Corinne, Danielle, Dave, Eda, Hanna, Henrik, Hsuan-Yu, Jiri, Peter, Samuel, and Vanessa.

I would not be at this point without the best and coolest parents in the world: Annette and Detlef. Thank you for letting me follow in your footsteps. A big thanks to my sister Vera for having been my best friend all her life. My final gratitude goes to Gidon – thank you for everything!

11. Curriculum Vitae

Lea Maria Bartsch

Knueslistrasse 10
8004 Zürich (Switzerland)

Mobile+ 41 (0) 76 699 0706
Email: l.bartsch@psychologie.uzh.ch
Birthday: 07/06/1992
Nationality: German

Professional Experience

10/2015 – 10/2018 **University of Zurich,** (Switzerland)
Department: **Cognition Lab and University Research Priority Program "Dynamics of Healthy Aging"**
Doctoral Student

- Realisation of behavioural experiments with young and older adults
- Realisation of fMRI experiments with young and older adults
- Programming of task-sets via Presentation® and Matlab
- Data Analysis and computational modelling with R
- Analysis of functional MRI-data with SPM, Freesurfer and the Princeton MVPA toolbox
- Writing peer reviewed research papers
- Presenting data at international conferences
- Teaching in the psychology master program

02/2015 – 05/2015 **Max-Planck-Institute for Human Development Berlin,** (Germany)
Department: **CONMEM Group, Jacobs-Study of Child Development**
Research Internship

- Recruitment of subjects
- Realisation of behavioural experiments with children
- Programming of task-sets via Presentation®
- Analysis of longitudinal structural MRI-data via Freesurfer and Matlab

03/2014 – 01/2015 **Technical University Dresden,** (Germany)
06/2015 – 07/2015 Department: **Lifespan Developmental Neuroscience**
Student Assistant Job

- Recruitment of subjects across all age groups
- Realisation of EEG experiments with old adults
- Realisation of TMS experiments with young adults
- Programming of task-sets via Presentation®

03/2013 – 05/2013 **Universitätsklinikum Carl Gustav Carus,** (Germany)

Department: Clinic and Polyclinic for Psychotherapy and Psychosomatic Medicine
Internship

- Observation of single and group therapy
- Conducting preliminary talks with patients
- Writing medical reports under supervision

01/2013 – 06/2012 **Technical University Dresden,** (Germany)
Department: Developmental Psychology
Research Internship
"Pop into Mind - How automatic is prospective memory? "

- Preparation, realisation and post-processing of Experiments with young and old adults
- Recruitment of subjects across both age groups
- Data analysis with SPSS and Eprime

Academic Career

09/2010 – 10/2015 **Technical University Dresden** (Germany)
Psychology: Cognitive Affective Neuroscience (M.Sc.)
Grade point average: 1.4 (very good)
Master Thesis: *Mimicking effects of aging: The role of the left dorsolateral PFC in model-based learning in younger adults*

10/2010 – 07/2013 **Technical University Dresden** (Germany)
Psychology (B.Sc.)
Grade point average: 1.9 (good)
Bachelor Thesis: *Pop Into Mind - How automatic is prospective memory?*

02/2008 – 07/2010 **Hermann-von-Helmholtz High School**, Potsdam (Germany)
Higher education entrance qualification
Grade point average: 1.5 (very good)

08/2007 – 02/2008 **École Secondaire de Mortagne**, Boucherville (Canada)
Exchange half year during High School

09/2002 – 08/2007 **Hermann-von-Helmholtz High School**, Potsdam (Germany)
Special Education Program (*Leistungsprofilklasse*), finished the academic lower cycle of higher general education in 5 instead of 6 years

01/1998 – 07/2002 V**aalserquatier Elementary School**, Aachen (Germany)

11/1997 – 01/1999 **Doyle Elementary School**, San Diego, California (USA)

Publications

Bartsch, L. M., Loaiza, V.M., Jäncke, L., Oberauer, K. & Lewis-Peacock, J.A. (2018). Unravelling refreshing and elaboration as separable cognitive processes in the brain with distinct impacts on working memory and long-term memory across the lifespan. *Unpublished Manuscript.*

Bartsch, L. M., Loaiza, V.M., & Oberauer, K. (2018). Does limited working-memory capacity underlie age differences in associative long-term memory? *Unpublished Manuscript.*

Bartsch, L. M., Singmann, H., & Oberauer, K. (2018). The effects of refreshing and elaboration on working memory performance, and their contributions to long-term memory formation. *Memory and Cognition*, 46(5), 1–13. https://doi.org/10.3758/s13421-018-0805-9

Wittkuhn, L., Eppinger, B., Bartsch, L. M., Thurm, F., Korb, F. M., & Li, S.-C. (2017). Repetitive transcranial magnetic stimulation over dorsolateral prefrontal cortex modulates value-based learning during sequential decision-making. *NeuroImage*, *167*. https://doi.org/10.1016/j.neuroimage.2017.11.057

Conference contributions

Bartsch, L.M. & Oberauer, K. (2017). The effect of refreshing and elaboration on working memory performance and their contribution to long-term memory formation. Talk at the 59th Conference of Experimental Psychologists (TeaP), Dresden, Germany.

Bartsch, L.M. & Oberauer, K. (2017). Elaboration in Working Memory and its Age-dependant Effect on Episodic Long-term Memory. Poster at the 4th International Conference Aging and Cognition and 2nd Conference of the European Cognitive Ageing Society, Zurich, Switzerland.

Bartsch, L.M. & Oberauer, K. (2017). The effect of refreshing and elaboration on working memory performance and their contribution to long-term memory formation. Talk at Workshop on The crossroads of attention in working memory, Hotel Bains D'Ovronnaz, Switzerland.

Bartsch, L.M. & Oberauer, K. (2017). The effect of refreshing and elaboration on working and long-term memory. Talk at the 20th Conference of the European Society for Cognitive Psychology (ESCoP), Potsdam, Germany.

Bartsch, L.M. & Oberauer, K. (2017). The Effect of Refreshing and Elaboration on Working and Long-term Memory. Poster at the Psychonomic Society 58th Annual Meeting, Vancouver, Canada.

Bartsch, L.M., Loaiza, V.M. & Oberauer, K. (2018) Importance of bindings in working memory to age differences in episodic long-term memory. Poster at the 3rd International Meeting of the Psychonomic Society, Amsterdam, The Netherlands.

Bartsch, L.M., Loaiza, V.M. & Oberauer, K. (2018) Importance of bindings in working memory to age differences in episodic long-term memory. Talk at the Ninth European Working Memory Symposium (EWOMS-9), Pavia, Italy.

Skills

IT	MS Office: Word, PowerPoint, Outlook and Excel; MATLAB, R, Freesurfer, FSL, SPM, Princeton MVPA toolbox; Presentation	
Languages	*German:*	Mother Tongue
	English:	Advanced User (in speech and writing)
	French:	Advanced User (in speech and writing)

12. References

Addis, D. R., & McAndrews, M. P. (2006). Prefrontal and hippocampal contributions to the generation and binding of semantic associations during successful encoding. *NeuroImage*, *33*(4), 1194–1206. https://doi.org/10.1016/j.neuroimage.2006.07.039

Amer, T., Giovanello, K. S., Grady, C. L., & Hasher, L. (2018). Age differences in memory for meaningful and arbitrary associations: A memory retrieval account. *Psychology and Aging*, *33*(1), 74.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In *Psychology of learning and motivation* (Vol. 2, pp. 89–195). Elsevier. https://doi.org/10.1016/S0079-7421(08)60422-3

Axmacher, N., Henseler, M. M., Jensen, O., Weinreich, I., Elger, C. E., & Fell, J. (2010). Cross-frequency coupling supports multi-item working memory in the human hippocampus. *Proceedings of the National Academy of Sciences*, *107*(7), 3228–3233. https://doi.org/10.1073/pnas.0911531107

Baddeley, A. D. (1986). *Working Memory*. Oxford: Oxford University Press.

Baddeley, A. D. (2012). Working memory: theories, models, and controversies. *Annual Review of Psychology*, *63*, 1–29.

Bailey, H., Dunlosky, J., & Hertzog, C. (2009). Does differential strategy use account for age-related deficits in working-memory performance? *Psychology and Aging*, *24*(1), 82–92. https://doi.org/10.1037/a0014078

Bailey, H., Dunlosky, J., & Kane, M. J. (2008). Why does working memory span predict complex cognition? Testing the strategy affordance hypothesis. *Memory & Cognition*, *36*(8), 1383–1390. https://doi.org/10.3758/MC.36.8.1383

Bailey, H., Dunlosky, J., & Kane, M. J. (2011). Contribution of strategy use to performance on complex and simple span tasks. Memory & Cognition, 39(3), 447–461. https://doi.org/10.3758/s13421-010-0034-3

Baker, J. T., Sanders, A. L., Maccotta, L., & Buckner, R. L. (2001). Neural correlates of verbal memory encoding during semantic and structural processing tasks. *Lippincott Williams & Wilkins*, *12*(8), 1251–1256. https://doi.org/10.1097/00001756-200105080-00039

Baltes, P. B., & Lindenberger, U. (1997). Emergence of a powerful connection between sensory and cognitive functions across the adult life span: a new window to the study of cognitive aging? *Psychology and Aging*, *12*(1), 12.

Barrouillet, P., & Camos, V. (2015). Working memory: Loss and reconstruction.

Barrouillet, P., Bernardin, S., & Camos, V. (2004). Time Constraints and Resource Sharing in Adults' Working Memory Spans. *Journal of Experimental Psychology: General*, *133*(1), 83–100. https://doi.org/10.1037/0096-3445.133.1.83

Barrouillet, P., Bernardin, S., & Camos, V. (2004). Time Constraints and Resource Sharing in Adults' Working Memory Spans. *Journal of Experimental Psychology: General*, *133*(1), 83–100. https://doi.org/10.1037/0096-3445.133.1.83

Barrouillet, P., Bernardin, S., Portrat, S., Vergauwe, E., & Camos, V. (2007). Time and cognitive load in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(3), 570–585. https://doi.org/10.1037/0278-7393.33.3.570

Bartsch, L. M., Singmann, H., & Oberauer, K. (2018). The effects of refreshing and elaboration on working memory performance , and their contributions to long-term memory formation. *Memory & Cognition*. https://doi.org/10.3758/s13421-018-0805-9

Bastin, C., & Van Der Linden, M. (2005). The effects of aging on the recognition of different types of associations. *Experimental Aging Research*, *32*(1), 61–77. https://doi.org/10.1080/03610730500326291

Batchelder, W. H., & Riefer, D. M. (1999). Theoretical and empirical review of multinomial process tree modeling. *Psychonomic Bulletin & Review*, *6*(1), 57–86.

Bayliss, D. M., Bogdanovs, J., & Jarrold, C. (2015). Consolidating working memory: Distinguishing the effects of consolidation, rehearsal and attentional refreshing in a working memory span task. *Journal of Memory and Language*, *81*, 34–50.

Blumenfeld, R. S. (2006). Dorsolateral Prefrontal Cortex Promotes Long-Term Memory Formation through Its Role in Working Memory Organization. *Journal of Neuroscience*, *26*(3), 916–925. https://doi.org/10.1523/JNEUROSCI.2353-05.2006

Blumenfeld, R. S., & Ranganath, C. (2007). Prefrontal cortex and long-term memory encoding: an integrative review of findings from neuropsychology and neuroimaging. *Neuroscientist*, *13*(3), 280–291. https://doi.org/10.1177/1073858407299290

Blumenfeld, R. S., Parks, C. M., & Yonelinas, A. P. (2010). Putting the Pieces Together : The

Role of Dorsolateral Prefrontal Cortex in Relational Memory Encoding, 257–265.

Borg, C., Leroy, N., Favre, E., Laurent, B., & Thomas-Antérion, C. (2011). How emotional pictures influence visuospatial binding in short-term memory in ageing and Alzheimer's disease? *Brain and Cognition*, *76*(1), 20–25. https://doi.org/10.1016/j.bandc.2011.03.008

Boujut, A., & Clarys, D. (2016). The effect of ageing on recollection: the role of the binding updating process. *Memory*, *24*(9), 1231–1242. https://doi.org/10.1080/09658211.2015.1091893

Brockmole, J. R., Parra, M. A., Sala, S. D., & Logie, R. H. (2008). Do binding deficits account for age-related decline in visual working memory? *Psychonomic Bulletin & Review*, *15*(3), 543–547. https://doi.org/10.3758/PBR.15.3.543

Brown, L. A., & Brockmole, J. R. (2010). The role of attention in binding visual features in working memory: Evidence from cognitive ageing. *Quarterly Journal of Experimental Psychology*, *63*(10), 2067–2079. https://doi.org/10.1080/17470211003721675

Brown, L. A., Niven, E. H., Logie, R. H., Rhodes, S., & Allen, R. J. (2017). Visual feature binding in younger and older adults: encoding and suffix interference effects. *Memory*, *25*(2), 261–275. https://doi.org/10.1080/09658211.2016.1156705

Buchler, N. G., Faunce, P., Light, L. L., Gottfredson, N., & Reder, L. M. (2011). Effects of repetition on associative recognition in young and older adults: Item and associative strengthening. *Psychology and Aging*, *26*(1), 111–126. https://doi.org/10.1037/a0020816

Buchner, A., Erdfelder, E., & Vaterrodt-Plünnecke, B. (1995). Toward unbiased measurement of conscious and unconscious memory processes within the process dissociation framework. *Journal of Experimental Psychology: General*, *124*(2), 137.

Cabeza, R., Anderson, N. D., Locantore, J. K., & McIntosh, A. R. (2002). Aging gracefully: compensatory brain activity in high-performing older adults. *Neuroimage*, *17*(3), 1394–1402.

Camos, V., & Portrat, S. (2015). The impact of cognitive load on delayed recall. *Psychonomic Bulletin & Review*, *22*(4), 1029–1034. https://doi.org/10.3758/s13423-014-0772-5

Camos, V., Johnson, M. R., Loaiza, V. M., Portrat, S., Souza, A. S., & Vergauwe, E. (2018). What is attentional refreshing in working memory? *Annals of the New York Academy of Sciences*, 1–14. https://doi.org/10.1111/nyas.13616

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., … Riddell,

A. (2017). *Stan* : A Probabilistic Programming Language. *Journal of Statistical Software*, *76*(1). https://doi.org/10.18637/jss.v076.i01

Chalfonte, B. L., & Johnson, M. K. (1996). Feature memory and binding in young and older adults. *Memory & Cognition*, *24*(4), 403–416. https://doi.org/10.3758/BF03200930

Chen, T., & Naveh-Benjamin, M. (2012). Assessing the associative deficit of older adults in long-term and short-term/working memory. *Psychology and Aging*, *27*(3), 666–682. https://doi.org/10.1037/a0026943

Christophel, T. B., Iamshchinina, P., Yan, C., Allefeld, C., & Haynes, J. D. (2018). Cortical specialization for attended versus unattended working memory. *Nature Neuroscience*, *21*(4), 494–496. https://doi.org/10.1038/s41593-018-0094-4

Chun, M. M., & Potter, M. C. (1995). A Two-Stage Model for Multiple Target Detection in Rapid Serial Visual Presentation. *Journal of Experimental Psychology*, *21*(1), 109–127.

Cohen, N. J., & Eichenbaum, H. (1993). *Memory, amnesisa and the hippocampal system*. MIT press.

Cooper, E., Greve, A., & Henson, R. N. (2017). Assumptions behind scoring source versus item memory: Effects of age, hippocampal lesions and mild memory problems. *Cortex*, *91*, 297–315. https://doi.org/10.1016/j.cortex.2017.01.001

Cowan, N. (1995). *Attention and memory: An integrated framework*. New York: NY: Oxford University Press.

Cowan, N. (2008). What are the differences between long-term, short-term, and working memory? *Progress in Brain Research*, *169*(07), 323–338. https://doi.org/10.1016/S0079-6123(07)00020-9.What

Cowan, N., Naveh-Benjamin, M., Kilb, A., & Saults, J. S. (2006). Life-span development of visual working memory: When is feature binding difficult? *Developmental Psychology*, *42*(6), 1089–1102. https://doi.org/10.1037/0012-1649.42.6.1089

Craik, F. I. M., & Byrd, M. (1982). Aging and cognitive deficits. In *Aging and cognitive processes* (pp. 191–211). Springer.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*(6), 671–684. https://doi.org/10.1016/S0022-5371(72)80001-X

Crowder, R. G. (1982). The demise of short-term memory. *Acta Psychologica*, *50*(3), 291–323.

D'Esposito, M., & Postle, B. R. (2015). The Cognitive Neuroscience of Working Memory, (66), 115–142. https://doi.org/10.1146/annurev-psych-010814-015031.THE

Davachi, L., Maril, a, & Wagner, a D. (2001). When keeping in mind supports later bringing to mind: neural markers of phonological rehearsal predict subsequent remembering. *Journal of Cognitive Neuroscience*, *13*(8), 1059–1070. https://doi.org/10.1162/089892901753294356

De Schrijver, S., & Barrouillet, P. (2017). Consolidation and restoration of memory traces in working memory. *Psychonomic Bulletin & Review*, *24*(5), 1651–1657.

De Simoni, C., & von Bastian, C. C. (2018). Working memory updating and binding training: Bayesian evidence supporting the absence of transfer, (April). https://doi.org/10.1037/xge0000453

Dunlosky, J., & Hertzog, C. (1998). Aging and deficits in associative memory: What is the role of strategy production? *Psychology and Aging*, *13*(4), 597.

Dunlosky, J., & Kane, M. J. (2007). The contributions of strategy use to working memory span: A comparison of strategy assessment methods. *The Quarterly Journal of Experimental Psychology*, *60*(9), 1227–1245. https://doi.org/10.1080/17470210600926075

Dunlosky, J., Hertzog, C., & Powell-Moman, A. (2005). The contribution of mediator-based deficiencies to age differences in associative learning. *Developmental Psychology*, *41*(2), 389.

Eriksson, J., Vogel, E. K., Lansner, A., Bergström, F., & Nyberg, L. (2015). Neurocognitive Architecture of Working Memory. *Neuron*, *88*(1), 33–46. https://doi.org/10.1016/j.neuron.2015.09.020

Fandakova, Y., Shing, Y. L., & Lindenberger, U. (2013). Differences in binding and monitoring mechanisms contribute to lifespan age differences in false memory. *Developmental Psychology*, *49*(10), 1822–32. https://doi.org/10.1037/a0031361

Fine, H. C., Shing, Y. L., & Naveh-Benjamin, M. (2018). Effects of changes in schematic support and of item repetition on age-related associative memory deficits: Theoretically-driven empirical attempts to reduce older adults' high false alarm rate. *Psychology and Aging*, *33*(1), 57.

Fletcher, P. C., Shallice, T., & Dolan, R. J. (2000). "Sculpting the response space" - An account of left prefrontal activation at encoding. *NeuroImage*, *12*(4), 404–417. https://doi.org/10.1006/nimg.2000.0633

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). “Mini-mental state”: a practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, *12*(3), 189–198. https://doi.org/10.1016/0022-3956(75)90026-6

Gallistel, C. R. (2009). The importance of proving the null. *Psychological Review*, *116*(2), 439–453. https://doi.org/10.1037/a0015251

Gallo, D. A., Meadow, N. G., Johnson, E. L., & Foster, K. T. (2008). Deep levels of processing elicit a distinctiveness heuristic: Evidence from the criterial recollection task. *Journal of Memory and Language*, *58*(4), 1095–1111. https://doi.org/10.1016/j.jml.2007.12.001

Gallo, D. A., Meadow, N. G., Johnson, E. L., & Foster, K. T. (2008). Deep levels of processing elicit a distinctiveness heuristic: Evidence from the criterial recollection task. Journal of Memory and Language, 58(4), 1095–1111. https://doi.org/10.1016/j.jml.2007.12.001

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). Bayesian data analysis, 3rd edition. Chapman & Hall/CRC.

Gilchrist, A. L., & Cowan, N. (2011). Can the focus of attention accommodate multiple, separate items? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(6), 1484.

Grady, C. L., McIntosh, A. R., Bookstein, F., Horwitz, B., Rapoport, S. I., & Haxby, J. V. (1998). Age-related changes in regional cerebral blood flow during working memory for faces. *Neuroimage*, *8*(4), 409–425.

Greene, R. L. (1987). Effects of maintenance rehearsal on human memory. *Psychological Bulletin*, *102*(3), 403–413. https://doi.org/10.1037/0033-2909.102.3.403

Hale, S., Rose, N. S., Myerson, J., Strube, M. J., Sommers, M., Tye-murray, N., & Spehar, B. (2011). The structure of working memory abilities across the adult life span. *Psychology and Aging*, *26*(1), 92–110. https://doi.org/10.1037/a0021483.The

Hannula, D. E., & Ranganath, C. (2008). Medial Temporal Lobe Activity Predicts Successful Relational Memory Binding. *Journal of Neuroscience*, *28*(1), 116–124. https://doi.org/10.1523/JNEUROSCI.3086-07.2008

Hara, Y., & Naveh-Benjamin, M. (2015). The role of reduced working memory storage and processing resources in the associative memory deficit of older adults: Simulation studies with younger adults. *Aging, Neuropsychology, and Cognition*, *22*(2), 129–154. https://doi.org/10.1080/13825585.2014.889650

Hartley, A. (2006). Changing role of the speed of processing construct in the cognitive

psychology of human aging. *Handbook of the Psychology of Aging*, 183–207.

Haxby, J. V, Connolly, A. C., & Guntupalli, J. S. (2014). Decoding Neural Representational Spaces Using Multivariate Pattern Analysis. *Annual Review of Neuroscience*, (June), 435–456. https://doi.org/10.1146/annurev-neuro-062012-170325

Haxby, J. V, Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., & Pietrini, P. (2001). Distrubuted and Overlapping Representations of Face and Objects in Ventral Temporal Cortex. *Science*, *293*(5539), 2425–2430. https://doi.org/10.1126/science.1063736

Haynes, J., & Rees, G. (2006). Decoding mental states from brain activity in humans. *Nature Reviews Neuroscience*, *7*(7), 523–534. https://doi.org/10.1038/nrn1931

Heck, D. W., Arnold, N. R., & Arnold, D. (2018). TreeBUGS: An R package for hierarchical multinomial-processing-tree modeling. *Behavior Research Methods*, *50*(1), 264–284. https://doi.org/10.3758/s13428-017-0869-7

Hoareau, V., Lemaire, B., Portrat, S., & Plancher, G. (2016). Reconciling Two Computational Models of Working Memory in Aging. *Topics in Cognitive Science*, *8*(1), 264–278. https://doi.org/10.1111/tops.12184

Hoyer, W. J., & Verhaeghen, P. (2006). *Memory aging*. *Handbook of the psychology of aging* (Vol. 6).

Jacoby, L. L. (1999). Ironic Effects of Repetition: Measuring Age-Related Differences in Memory. *Journal of Experimental Psychology: Learning Memory and Cognition*, *25*(1), 3–22. https://doi.org/10.1037/0278-7393.25.1.3

Jacoby, L. L., Debner, J. A., & Hay, J. F. (2001). Proactive Interference, Accessibility Bias, and Process Dissociations: Valid Subjective Reports of Memory. *Journal of Experimental Psychology: Learning Memory and Cognition*, *27*(3), 686–700. https://doi.org/10.1037/0278-7393.27.3.686

Jaswal, S., & Logie, R. H. (2011). Configural encoding in visual feature binding. *Journal of Cognitive Psychology*, *23*(5), 586–603. https://doi.org/10.1080/20445911.2011.570256

Johnson, M. K. (1992). MEM : Mechanisms of Recollection. *Journal of Cognitive Neuroscience*, *4*(3), 268–280.

Johnson, M. K., Mitchell, K. J., Raye, C. L., & Greene, E. J. (2004). An age-related deficit in prefrontal cortical function associated with refreshing information. *Psychological Science*, *15*(2), 127–132. https://doi.org/10.1016/j.neuroscience.2008.09.014

Johnson, M. K., Raye, C. L., Mitchell, K. J., Greene, E. J., & Adam, W. (2003). fMRI Evidence for an Organization of Prefrontal Cortex by Both Type of Proces ..., 265–273.

Johnson, M. K., Raye, C. L., Mitchell, K. J., Greene, E. J., Cunningham, W. a, & Sanislow, C. a. (2005). Using fMRI to investigate a component process of reflection: prefrontal correlates of refreshing a just-activated representation. *Cognitive, Affective & Behavioral Neuroscience*, *5*(3), 339–361. https://doi.org/10.3758/CABN.5.3.339

Johnson, M. K., Reeder, J. a, Raye, C. L., & Mitchell, K. J. (2002). Second thoughts versus second looks: an age-related deficit in reflectively refreshing just-activated information. *Psychological Science : A Journal of the American Psychological Society / APS*, *13*(1), 64–67. https://doi.org/10.1111/1467-9280.00411

Johnson, M. R., & Johnson, M. K. (2009). Toward characterizing the neural correlates of component processes of cognition. *Neuroimaging of Human Memory: Linking Cognitive Processes to Neural Systems*, 169–194.

Jolicœur, P., & Dell'Acqua, R. (1998). The Demonstration of Short-Term Consolidation. *Cognitive Psychology*, *36*(2), 138–202. https://doi.org/10.1006/cogp.1998.0684

Jonides, J., Marshuetz, C., Smith, E. E., Reuter-Lorenz, P. A., Koeppe, R. A., & Hartley, A. (2000). Age differences in behavior and PET activation reveal differences in interference resolution in verbal working memory. *Journal of Cognitive Neuroscience*, *12*(1), 188–196.

Kaakinen, J. K., & Hyönä, J. (2007). Strategy use in the reading span test: An analysis of eye movements and reported encoding strategies. *Memory*, *15*(6), 634–646. https://doi.org/10.1080/09658210701457096

Kapur, S., Craik, F. I., Tulving, E., Wilson, A. A., Houle, S., & Brown, G. M. (1994). Neuroanatomical correlates of encoding in episodic memory: levels of processing effect. *Proceedings of the National Academy of Sciences of the United States of America*, *91*(6), 2008–2011. https://doi.org/10.1073/pnas.91.6.2008

Kausler, D. H. (1994). *Learning and memory in normal aging.* Academic Press.

Kim, S. Y., & Giovanello, K. S. (2011). The Effects of Attention on Age-Related Relational Memory Deficits: Evidence From a Novel Attentional Manipulation. *Psychology and Aging*, *26*(3), 678–688. https://doi.org/10.1037/a0022326

Klatsky, R. (1988). Theories of information processing and theories of aging. In L. Light & D. M. Burke (Eds.), *Language, memory, and aging* (pp. 1–16). New York: Cambridge University

Press.

Kriegeskorte, N., & Bandettini, P. A. (2007). Analyzing for information, not activation, to exploit high-resolution fMRI. *NeuroImage*, *38*(4), 649–662. https://doi.org/10.1016/j.neuroimage.2007.02.022

Kriegeskorte, N., Simmons, W., Bellgowan, P., & Baker, C. (2009). Circular analysis in systems neuroscience – the dangers of double dipping. *Nat Neurosci*, *12*(5), 535–540. https://doi.org/10.1038/nn.2303.Circular

LaRocque, J. J., Riggall, A. C., Emrich, S. M., & Postle, B. R. (2017). Within-Category Decoding of Information in Different Attentional States in Short-Term Memory. *Cerebral Cortex*, *27*(10), 4881–4890. https://doi.org/10.1093/cercor/bhw283

Lehrl, S. (2005). *Mehrfachwahl-Wortschatz-Intelligenztest MWT-B* (5th ed.). Balingen: Spitta Verlag.

Lemaire, B., Pageot, A., Plancher, G., & Portrat, S. (2017). What is the time course of working memory attentional refreshing? *Psychonomic Bulletin & Review*. https://doi.org/10.3758/s13423-017-1282-z

Lewandowsky, S., & Farrell, S. (2008). Short-term memory: New data and a model. *Psychology of Learning and Motivation*, *49*, 1–48.

Lewis-Peacock, J. A., & Norman, K. A. (2013). Multi-voxel pattern analysis of fMRI data. *The Cognitive Neurosciences V*, (512). Retrieved from http://compmem.princeton.edu/pmwiki/uploads/Lewis_Peacock_Norman_CogNeuro.pdf%5Cnpapers3://publication/uuid/6557C674-037D-46BA-BA70-3274007197F0

Lewis-Peacock, J. A., & Norman, K. A. (2014). Multi-Voxel Pattern Analysis of fMRI Data. In M. S. Gazzaniga & G. R. Mangun (Eds.), *The Cognitive Neurosciences* (5th ed., pp. 911–920). Cambridge, Massachusetts: MIT Press.

Lewis-Peacock, J. A., & Postle, B. R. (2008). Temporary Activation of Long-Term Memory Supports Working Memory, *28*(35), 8765–8771. https://doi.org/10.1523/JNEUROSCI.1953-08.2008.Temporary

Lewis-Peacock, J. A., & Postle, B. R. (2012). Decoding the internal focus of attention. *Neuropsychologia*, *50*(4), 470–478. https://doi.org/10.1016/j.neuropsychologia.2011.11.006

Lewis-Peacock, J. A., Drysdale, A. T., & Postle, B. R. (2015). Neural evidence for the flexible control of mental representations. *Cerebral Cortex*, *25*(10), 3303–3313.

https://doi.org/10.1093/cercor/bhu130

Lewis-Peacock, J. A., Drysdale, A. T., Oberauer, K., & Postle, B. R. (2012). Neural evidence for a distinction between short-term memory and the focus of attention. *Journal of Cognitive Neuroscience*, *24*(1), 61–79. https://doi.org/10.1162/jocn_a_00140

Li, S.-C., Lindenberger, U., & Sikström, S. (2001). Aging cognition: from neuromodulation to representation. *Trends in Cognitive Sciences*, *5*(11), 479–486.

Li, S.-C., Lindenberger, U., Hommel, B., Aschersleben, G., Prinz, W., & Baltes, P. B. (2004). Transformations in the couplings among intellectual cognitive processes across the life span. *Psychological Science*, *15*(3), 155–163. https://doi.org/10.1111/j.0956-7976.2004.01503003.x

Loaiza, V. M., & Camos, V. (2016). Does Controlling for Temporal Parameters Change the Levels-of-Processing Effect in Working Memory? *Advances in Cognitive Psychology*, *12*(1), 2–9. https://doi.org/10.5709/acp-0182-3

Loaiza, V. M., & Mccabe, D. P. (2012). Temporal–contextual processing in working memory: Evidence from delayed cued recall and delayed free recall tests. *Memory & Cognition*, *40*(2), 191–203. https://doi.org/10.3758/s13421-011-0148-2

Loaiza, V. M., McCabe, D. P., Youngblood, J. L., Rose, N. S., & Myerson, J. (2011). The influence of levels of processing on recall from working memory and delayed recall tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(5), 1258.

Loftus, G. R. (1978). On interpretation of interactions. *Memory & Cognition*, *6*(3), 312–319. https://doi.org/10.3758/BF03197461

Maillet, D., & Rajah, M. N. (2014). Age-related differences in brain activity in the subsequent memory paradigm: A meta-analysis. *Neuroscience and Biobehavioral Reviews*, *45*, 246–257. https://doi.org/10.1016/j.neubiorev.2014.06.006

McElree, B. (1998). Attended and Non-Attended States in Working Memory: Accessing Categorized Structures. *Journal of Memory and Language*, *38*(2), 225–252. https://doi.org/10.1006/jmla.1997.2545

McElree, B. (2006). Accessing recent events. *Psychology of Learning and Motivation*, *46*, 155–200.

Melton, A. W. (1963). Implications of short-term memory for a general theory of memory. *Journal of Verbal Learning and Verbal Behavior*, *2*(1), 1–21.

Mitchell, K. J., Johnson, M. K., Raye, C. L., Mather, M., & Esposito, M. D. (2000). Aging and Reflective Processes of Working Memory: Binding and Test Load Deficits. *Psychology and Aging*, *15*(3), 527–541. https://doi.org/10.1037/70882-

Morey, R. D., Hoekstra, R., Rouder, J. N., Lee, M. D., & Wagenmakers, E.-J. (2016). The fallacy of placing confidence in confidence intervals. *Psychonomic Bulletin & Review*, *23*(1), 103–123. https://doi.org/10.3758/s13423-015-0947-8

Morey, R. D., Rouder, J. N., & Jamil, T. (2015). BayesFactor: Computation of Bayes factors for common designs. *R Package Version 0.9*, *9*.

Murray, L. J., & Ranganath, C. (2007). The dorsolateral prefrontal cortex contributes to successful relational memory encoding. *J Neurosci*, *27*(20), 5515–5522. https://doi.org/10.1523/JNEUROSCI.0406-07.2007

Nairne, J. S. (1990). A feature model of immediate memory. *Memory & Cognition*, *18*(3), 251–269.

Nairne, J. S. (2002). Remembering over the short-term: The case against the standard model. *Annual Review of Psychology*, *53*(1), 53–81.

Naveh-Benjamin, M. (2000). Adult age differences in memory performance: Tests of an associative deficit hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(5), 1170–1187. https://doi.org/10.1037/0278-7393.26.5.1170

Naveh-Benjamin, M., & Old, S. R. (2008). 2.41 Aging and Memory. *Learning and Memory: A Comprehensive Reference*, 787–808. https://doi.org/http://dx.doi.org/10.1016/B978-012370509-9.00133-9

Noack, H., Lövdén, M., & Lindenberger, U. (2012). Normal aging increases discriminal dispersion in visuospatial short-term memory. *Psychology and Aging*, *27*(3), 627–637. https://doi.org/10.1037/a0027251

Norman, K. A., Polyn, S. M., Detre, G. J., & Haxby, J. V. (2006). Beyond mind-reading: multi-voxel pattern analysis of fMRI data. *Trends in Cognitive Sciences*, *10*(9), 424–430. https://doi.org/10.1016/j.tics.2006.07.005

Oberauer, K. (2002). Access to information in working memory: exploring the focus of attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(3), 411.

Oberauer, K. (2005). Binding and Inhibition in Working Memory: Individual and Age Differences in Short-Term Recognition. *Journal of Experimental Psychology: General*,

134(3), 368–387. https://doi.org/10.1037/0096-3445.134.3.368

Oberauer, K. (2009). Design for a working memory. *Psychology of Learning and Motivation*, *51*, 45–100.

Oberauer, K., & Bialkova, S. (2009). Accessing information in working memory: Can the focus of attention grasp two elements at the same time? *Journal of Experimental Psychology: General*, *138*(1), 64.

Oberauer, K., & Hein, L. (2012). Attention to information in working memory. *Current Directions in Psychological Science*, *21*(3), 164–169.

Oberauer, K., & Kliegl, R. (2001). Beyond resources: Formal models of complexity effects and age differences in working memory. *European Journal of Cognitive Psychology*, *13*(1-2), 187–215. https://doi.org/10.1080/09541440042000278

Oberauer, K., & Lewandowsky, S. (2011). Modeling working memory: A computational implementation of the Time-Based Resource-Sharing theory. *Psychonomic Bulletin & Review*, *18*(1), 10–45.

Old, S. R., & Naveh-Benjamin, M. (2008). Differential effects of age on item and associative measures of memory: A meta-analysis. *Psychology and Aging*, *23*(1), 104–118. https://doi.org/10.1037/0882-7974.23.1.104

Overman, A. A., McCormick-Huhn, J. M., Dennis, N. A., Salerno, J. M., & Giglio, A. P. (2018). Older adults' associative memory is modified by manner of presentation at encoding and retrieval. *Psychology and Aging*, *33*(1), 82–92. https://doi.org/10.1037/pag0000215

Park, D. C., Lautenschlager, G., Hedden, T., Davidson, N. S., Smith, A. D., & Smith, P. K. (2002). Models of visuospatial and verbal memory across the adult life span. *Psychology and Aging*, *17*(2), 299–320. https://doi.org/10.1037/0882-7974.17.2.299

Parra, M. A., Abrahams, S., Logie, R. H., & Sala, S. Della. (2009). Age and binding within-dimension features in visual short-term memory. *Neuroscience Letters*, *449*(1), 1–5. https://doi.org/10.1016/j.neulet.2008.10.069

Penny, W. D., Friston, K. J., Ashburner, J. T., Kiebel, S. J., & Nichols, T. E. (2011). *Statistical parametric mapping: the analysis of functional brain images*. Elsevier.

Peterson, D. J., Schmidt, N. E., & Naveh-Benjamin, M. (2017). The role of schematic support in age-related associative deficits in short-term and long-term memory. *Journal of Memory and Language*, *92*, 79–97. https://doi.org/10.1016/j.jml.2016.05.007

Plancher, G., Boyer, H., Lemaire, B., & Portrat, S. (2017). Under Which Conditions Can Older Participants Maintain Information In Working Memory? *Experimental Aging Research*, *43*(5), 409–429. https://doi.org/10.1080/0361073X.2017.1369730

Portrat, S., & Lemaire, B. (2015). Is Attentional Refreshing in Working Memory Sequential? A Computational Modeling Approach. *Cognitive Computation*, *7*(3), 333–345. https://doi.org/10.1007/s12559-014-9294-8

Postle, B. R., Awh, E., Serences, J. T., Sutterer, D. W., & D'Esposito, M. (2013). The positional-specificity effect reveals a passive-trace contribution to visual short-term memory. *PLoS ONE*, *8*(12). https://doi.org/10.1371/journal.pone.0083483

R Core Team. (2017). R: A language and environment for statistical computing. Vienna, Austria; 2014. Retrieved from https://www.r-project.org/.

Ragland, J. D., Blumenfeld, R. S., Ramsay, I. S., Yonelinas, A. P., Yoon, J., Solomon, M., … Ranganath, C. (2012). Neural correlates of relational and item-specific encoding during working and long-term memory in schizophrenia. *NeuroImage*, *59*(2), 1719–1726. https://doi.org/10.1016/j.neuroimage.2011.08.055

Ranganath, C. (2006). Working memory for visual objects: Complementary roles of inferior temporal, medial temporal, and prefrontal cortex. *Neuroscience*, *139*(1), 277–289. https://doi.org/10.1016/j.neuroscience.2005.06.092

Ranganath, C., & Blumenfeld, R. S. (2005). Doubts about double dissociations between short- and long-term memory. *Trends in Cognitive Sciences*, *9*(8), 374–380. https://doi.org/10.1016/j.tics.2005.06.009

Ranganath, C., Cohen, M. M. X., & Brozinsky, C. C. J. (2005). Working memory maintenance contributes to long-term memory formation: neural and behavioral evidence. *Journal of Cognitive Neuroscience*, *17*(7), 994–1010. https://doi.org/10.1162/0898929054475118

Rankin, J. L., & Collins, M. (1985). Adult age differences in memory elaboration. *Journal of Gerontology*, *40*(4), 451–458.

Raye, C. L., Johnson, M. K., Mitchell, K. J., Greene, E. J., & Johnson, M. R. (2005). Refreshing: A Minimal Executive Function. *Control*, *43*(1), 135–145. https://doi.org/10.1016/S0010-9452(08)70451-9

Raye, C. L., Johnson, M. K., Mitchell, K. J., Greene, E. J., & Johnson, M. R. (2007). Refreshing : a Minimal Executive Function. *Cortex*, *43*(1), 135–145.

https://doi.org/10.1016/S0010-9452(08)70451-9

Raye, C. L., Johnson, M. K., Mitchell, K. J., Reeder, J. a, & Greene, E. J. (2002). Neuroimaging a single thought: dorsolateral PFC activity associated with refreshing just-activated information. *NeuroImage*, *15*(2), 447–453. https://doi.org/10.1006/nimg.2001.0983

Raye, C. L., Mitchell, K. J., Reeder, J. a, Greene, E. J., & Johnson, M. K. (2008). Refreshing one of several active representations: behavioral and functional magnetic resonance imaging differences between young and older adults. *Journal of Cognitive Neuroscience*, *20*(5), 852–862. https://doi.org/10.1162/jocn.2008.20508

Read, C. A., Rogers, J. M., & Wilson, P. H. (2016). Working memory binding of visual object features in older adults. *Aging, Neuropsychology, and Cognition*, *23*(3), 263–281. https://doi.org/10.1080/13825585.2015.1083937

Reuter-Lorenz, P. A., Jonides, J., Smith, E. E., Hartley, A., Miller, A., Marshuetz, C., & Koeppe, R. A. (2000). Age differences in the frontal lateralization of verbal and spatial working memory revealed by PET. *Journal of Cognitive Neuroscience*, *12*(1), 174–187.

Rhodes, S., Parra, M. A., & Logie, R. H. (2016). Ageing and feature binding in visual working memory: The role of presentation time. *Quarterly Journal of Experimental Psychology*, *69*(4), 654–668. https://doi.org/10.1080/17470218.2015.1038571

Ricker, T. J. (2015). The role of short-term consolidation in memory persistence. *AIMS Neurosci*, *2*, 259–279.

Ricker, T. J., & Cowan, N. (2014). Differences between presentation methods in working memory procedures: A matter of working memory consolidation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(2), 417.

Rose, N. S., & Craik, F. I. M. (2012). A processing approach to the working memory/long-term memory distinction: Evidence from the levels-of-processing span task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*(4), 1019–1029. https://doi.org/10.1037/a0026976

Rose, N. S., Buchsbaum, B. R., & Craik, F. I. M. (2014). Short-term retention of a single word relies on retrieval from long-term memory when both rehearsal and refreshing are disrupted. *Memory & Cognition*, *42*(5), 689–700. https://doi.org/10.3758/s13421-014-0398-x

Rose, N. S., Craik, F. I. M., & Buchsbaum, B. R. (2014). Levels of Processing in Working Memory: Differential Involvement of Fronto-Temporal Networks. *Journal of Cognitive*

Neuroscience, *27*(3), 522–532. https://doi.org/DOI: 10.1162/jocn_a_00738

Rose, N. S., LaRocque, J. J., Riggall, A. C., Gosseries, O., Starrett, M. J., Meyering, E. E., & Postle, B. R. (2016). Reactivation of latent working memories with transcranial magnetic stimulation. *Science*, *354*(6316), 1136 LP-1139. Retrieved from http://science.sciencemag.org/content/354/6316/1136.abstract

Rose, N. S., Myerson, J., Roediger III, H. L., Hale, S., Roediger, H. L., & Hale, S. (2010). Similarities and differences between working memory and long-term memory: Evidence from the levels-of-processing span task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(2), 471. https://doi.org/10.1037/a0018405

Rouder, J. N. (2014). Optional stopping: No problem for Bayesians. *Psychonomic Bulletin & Review*, *21*(2), 301–308.

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review*, *16*(2), 225–237. https://doi.org/10.3758/PBR.16.2.225

Rypma, B., & D'Esposito, M. (2000). Isolating the neural mechanisms of age-related changes in human working memory. *Nature Neuroscience*, *3*(5), 509.

Salthouse, T. A. (1996). The Processing-Speed Theory of Adult Age Diff erences in Cognition. *Psychological Review*, *103*(3), 403–428. https://doi.org/10.1037/0033-295X.103.3.403

Sander, M. C., Werkle-Bergner, M., & Lindenberger, U. (2011). Binding and strategic selection in working memory: A lifespan dissociation. *Psychology and Aging*, *26*(3), 612–624. https://doi.org/10.1037/a0023055

Schönbrodt, F. D., Wagenmakers, E.-J., Zehetleitner, M., & Perugini, M. (2017). Sequential hypothesis testing with Bayes factors: Efficiently testing mean differences. *Psychological Methods*, *22*(2), 322.

Shing, Y. L., Werkle-Bergner, M., Brehmer, Y., Müller, V., Li, S. C., & Lindenberger, U. (2010). Episodic memory across the lifespan: The contributions of associative and strategic components. *Neuroscience and Biobehavioral Reviews*, *34*(7), 1080–1091. https://doi.org/10.1016/j.neubiorev.2009.11.002

Shing, Y. L., Werkle-Bergner, M., Li, S.-C., & Lindenberger, U. (2008). Associative and strategic components of episodic memory: a life-span dissociation. *Journal of Experimental Psychology. General*, *137*(3), 495–513. https://doi.org/10.1037/0096-3445.137.3.495

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, *22*(11), 1359–1366.

Smith, A. D. (1980). Age differences in encoding, storage, and retrieval. *New Directions in Memory and Aging*, 23–46.

Smith, J. B., & Batchelder, W. H. (2010). Beta-MPT: Multinomial processing tree models for addressing individual differences. *Journal of Mathematical Psychology*, *54*(1), 167–183.

Souza, A. S., & Oberauer, K. (2017). The contributions of visual and central attention to visual working memory. *Attention, Perception, & Psychophysics*. https://doi.org/10.3758/s13414-017-1357-y

Souza, A. S., Vergauwe, E., & Oberauer, K. (2018). Where to attend next: guiding refreshing of visual, spatial, and verbal representations in working memory. *Annals of the New York Academy of Sciences*, 20–22. https://doi.org/10.1111/nyas.13621

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & van der Linde, A. (2002). Bayesian Measures of Model Complexity and Fit. *Journal of the Royal Statistical Society*, *64*(4), 583–639.

Stan Development Team. (2018). rstanarm: {Bayesian} applied regression modeling via {Stan}. Retrieved from http://mc-stan.org/

Treisman, A., & Zhang, W. (2006). Location and binding in visual working memory. *Memory & Cognition*, *34*(8), 1704–1719. https://doi.org/10.3758/BF03195932

Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organization of Memory* (pp. 381–403). New York: New York: Academic Press.

Vergauwe, E., & Langerock, N. (2017). Attentional refreshing of information in working memory: Increased immediate accessibility of just-refreshed representations. *Journal of Memory and Language*, *96*, 23–35.

Vergauwe, E., Camos, V., & Barrouillet, P. (2014). The impact of storage on processing: How is information maintained in working memory? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(4), 1072.

Wagenmakers, E.-J. (2007). A practical solution to the pervasive problems of p values. *Psychonomic Bulletin & Review*, *14*(5), 779–804.

Wagner, A. D., Poldrack, R. A., Eldridge, L. L., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. E. (1998). Lateralization of Prefrontal Activation During Episodic. *NeuroReport*, *9*(16),

3711–3717.

Wechsler, W. (1982). Manual for the Hamburg-Wechsler intelligence test for adults (HAWIE). *Bern, Switzerland: Huber*.

Wilhelm, O., Hildebrandt, A., & Oberauer, K. (2013). What is working memory capacity, and how can we measure it? *Frontiers in Psychology*, *4*(JUL), 1–22. https://doi.org/10.3389/fpsyg.2013.00433

Wixted, J. T. (2004). The Psychology and Neuroscience of Forgetting. *Annual Review of Psychology*, *55*(1), 235–269. https://doi.org/10.1146/annurev.psych.55.090902.141555

Zacks, R. T., & Hasher, L. (1994). Directed ignoring: Inhibitory regulation of working memory. In D. Dagenbach & T. H. Carr (Eds.), *Inhibitory processes in attention, memory, and language* (pp. pp. 241–264). San Diego, CA: CA: Academic Press.

Zacks, R. T., Hasher, L., & Li, K. Z. H. (2000). Human Memory. *The Handbook of Aging and Cognition*, 293–357. https://doi.org/10.3758/s13423-011-0086-9